Nigel Slack and Alistair Brandon-Jones

Operations and Process Management

Principles and Practice for Strategic Impact

Sixth Edition

Pearson

OPERATIONS AND PROCESS MANAGEMENT

Pearson

At Pearson, we have a simple mission: to help people make more of their lives through learning.

We combine innovative learning technology with trusted content and educational expertise to provide engaging and effective learning experiences that serve people wherever and whenever they are learning.

From classroom to boardroom, our curriculum materials, digital learning tools and testing programmes help to educate millions of people worldwide – more than any other private enterprise.

Every day our work helps learning flourish, and wherever learning flourishes, so do people.

To learn more, please visit us at **www.pearson.com/uk**

OPERATIONS AND PROCESS MANAGEMENT

Principles and Practice for Strategic Impact

Sixth Edition

Nigel Slack

Alistair Brandon-Jones

Pearson

Harlow, England • London • New York • Boston • San Francisco • Toronto • Sydney • Dubai • Singapore • Hong Kong
Tokyo • Seoul • Taipei • New Delhi • Cape Town • São Paulo • Mexico City • Madrid • Amsterdam • Munich • Paris • Milan

PEARSON EDUCATION LIMITED
KAO Two
KAO Park
Harlow
CM17 9NA
United Kingdom
Tel: +44 (0)1279 623623
Web: www.pearson.com/uk

First published 2006 (print)
Second edition published 2009 (print)
Third edition published 2012 (print and electronic)
Fourth edition published 2015 (print and electronic)
Fifth edition published 2018 (print and electronic)
Sixth edition published 2021 (print and electronic)

© Pearson Education Limited 2006, 2009 (print)
© Pearson Education Limited 2012, 2015, 2018, 2021 (print and electronic)

The rights of Nigel Slack and Alistair Brandon-Jones to be identified as authors of this work has been asserted by them in accordance with the Copyright, Designs and Patents Act 1988.

The print publication is protected by copyright. Prior to any prohibited reproduction, storage in a retrieval system, distribution or transmission in any form or by any means, electronic, mechanical, recording or otherwise, permission should be obtained from the publisher or, where applicable, a licence permitting restricted copying in the United Kingdom should be obtained from the Copyright Licensing Agency Ltd, Barnard's Inn, 86 Fetter Lane, London EC4A 1EN.

The ePublication is protected by copyright and must not be copied, reproduced, transferred, distributed, leased, licensed or publicly performed or used in any way except as specifically permitted in writing by the publishers, as allowed under the terms and conditions under which it was purchased, or as strictly permitted by applicable copyright law. Any unauthorised distribution or use of this text may be a direct infringement of the authors' and the publisher's rights and those responsible may be liable in law accordingly.

All trademarks used herein are the property of their respective owners. The use of any trademark in this text does not vest in the author or publisher any trademark ownership rights in such trademarks, nor does the use of such trademarks imply any affiliation with or endorsement of this book by such owners.

Pearson Education is not responsible for the content of third-party internet sites.

ISBN: 978-1-292-35006-6 (print)
978-1-292-35010-3 (PDF)
978-1-292-35011-0 (ePub)

British Library Cataloguing-in-Publication Data
A catalogue record for the print edition is available from the British Library

Library of Congress Cataloging-in-Publication Data
Names: Slack, Nigel, author. | Brandon-Jones, Alistair, author.
Title: Operations and process management : principles and practice for strategic impact / Nigel Slack, Alistair Brandon-Jones.
Description: Sixth edition. | Harlow, England ; London ; New York : Pearson, 2021. | Includes bibliographical references and index. | Description based on print version record and CIP data provided by publisher; resource not viewed.
Identifiers: LCCN 2020055631 (print) | LCCN 2020055632 (ebook) | ISBN 9781292350103 (PDF) | ISBN 9781292350110 (ePub) | ISBN 9781292350066 (print)
Subjects: LCSH: Manufacturing processes. | Process control. | Operations research.
Classification: LCC TS183 (ebook) | LCC TS183 .O64 2021 (print) | DDC 658.5–dc23
LC record available at https://lccn.loc.gov/2020055631
LC record available at https://lccn.loc.gov/2020055632

10 9 8 7 6 5 4 3 2 1
25 24 23 22 21

Front cover image: Jackal Pan/Moment/Getty Images
Cover design by Design Deluxe
Print edition typeset in Syntax Com of 9.25/12.5 pt Regular by SPi Global
Printed in Slovakia by Neografia

NOTE THAT ANY PAGE CROSS REFERENCES REFER TO THE PRINT EDITION

Brief contents

Contents

Companion Website

For open-access **student resources** specifically written to compliment this textbook and support your learning please visit **go.pearson.com/uk/he/resources**

Lecturer Resources

For password-protected online resources tailored to support use of this textbook in teaching please visit **go.pearson.com/uk/he/resources**

Guide to case studies

Chapter	Case name and description	Region	Manufac-turing/ service	Company size	Topics/techniques
Chapter 1 Operations and processes	Kaston-Trenton Service (KTS)	UK	M S	Medium	Role of operations, process objectives, types of operation and process
Chapter 2 Operations and strategic impact	IKEA looks to the future	World	S	Large	Operations strategy, operations objectives, strategic fit
Chapter 3 Product and service innovation	Widescale Studios and the Fierybryde development	UK	S	Small	Product development, operations strategy, process performance
Chapter 4 Operations scope and structure	Aarens Electronic	Europe	S/M	Medium	Location, capacity, scope of operations
Chapter 5 Process design 1 – Positioning	McPherson Charles Solicitors	UK	S	Medium	Process design, job design, process technology, process resourcing
Chapter 6 Process design 2 – Analysis	The Action Response Applications Processing Unit (ARAPU)	Africa, Asia, UK	S	Small	Process design, process mapping, balancing, Little's Law
Chapter 7 Supply chain management	Big or small? EDF's sourcing dilemma	UK	S M	Large (EDF); Small (local SMEs)	Sourcing strategy, supplier selection, supplier development
Chapter 8 Capacity management	FreshLunch	UK	S	Small	Demand forecasting and capacity planning
Chapter 9 Inventory management	supplies4medics.com	Europe	S	Medium	Inventory management, Inventory information systems, ABC analysis
Chapter 10 Resource planning and control	Audall Auto Servicing	UK	S	Medium	Planning and control, Gantt charts, activity monitoring, controlling activities
Chapter 11 Lean synchronisation	St Bridget's Hospital	Europe	S	Medium	Improvement, quality, application of lean principles
Chapter 12 Improvement	Ferndale Sands Conference Centre	Australia	S	Small	Improvement, performance, prioritisation
Chapter 13 Quality management	Rapposcience Labs	Netherlands	M	Medium	Improvement principles, statistical process control, process learning, operations capabilities
Chapter 14 Risk and resilience	Slagelse Industrial Services	Denmark	S M	Large	Risk, failure prevention, supplier selection, relationship management
Chapter 15 Project management	Kloud BV and Sakura Bank K.K.	Netherlands and Japan	S	Small (Kloud BV); Large (Sakura Bank K.K.)	Project planning (timing, costing, resourcing) and project risk

Preface

Why is operations and process management essential?

Because making operations and processes better will make the whole organisation better. Because operations and process management is about getting things done. Because without effective operations and processes there can be no long-term success for any organisation. Because it is at the heart of what all organisations do; they create value through their productive resources. Because it is the essential link that connects broad long-term strategy and day-to-day ongoing activities. This is why operations and process management has been changing. It has always been exciting, and it has always been challenging, but now it has acquired a much more prominent profile. The current edition reflects this in a number of ways.

This text stresses the importance of operations and process management

Of course, it has always been important, but increasingly managers in all types of enterprise are accepting that operations management can make or break their businesses. Effective operations management can keep costs down, enhance the potential to improve revenue, promote an appropriate allocation of capital resources and, most importantly, develop the capabilities that provide future competitive advantage.

This text stresses the real strategic impact of operations and process management

Operations are not always operational. The operations function also has a vital strategic dimension, and operations management is now expected to play a part in shaping strategic direction, not just responding to it.

This text stresses that operations and process management matters to all sectors of the economy

At one time, operations management was seen as being most relevant to manufacturing and a few types of mass-service businesses. Now the lessons are seen applying to all types of enterprise: all types of service and manufacturing, large or small organisations, public or private, for-profit or not-for-profit.

This text stresses that operations and process management is of interest to all managers

Perhaps most importantly, because operations management is accepted as being founded on the idea of managing process, and because managers in all functions of the business are now accepting that they spend much of their time managing processes, it is clear that, to some extent, all managers are operations managers. The principles and practice of operations management are therefore relevant to every manager.

This text extends the scope of operations and process management

The obvious unit of analysis of operations management is the operations function itself – the collection of resources and processes that produce products and services. But, if managers from other functions are to be included, operations management must also address itself to process management at a more generic level. Also, no operation can consider itself in isolation from its customers, suppliers, collaborators and competitors. It must see itself as part of the extended supply network. Operations management increasingly needs to work at all three levels of analysis – the individual process, the operation itself and the supply network.

All this has implications for the way operations management is studied, especially at post-experience and postgraduate levels, and the way operations management is practised. It has also very much shaped the way this text has been structured. In addition to covering all the important topics that make the subject so powerful, it places particular emphasis on the following:

- *Principles* – that is, the core ideas that describe how operations behave, how they can be managed and how they can be improved. These are not immutable laws or prescriptions that dictate how operations *should* be managed, nor are they descriptions that simply explain or categorise issues. But they are indications of important underlying ideas.
- *Diagnosis* – an approach that questions and explores the fundamental drivers of operations performance. Aims to uncover or 'diagnose' the underlying trade-offs that operations need to overcome, and the implications and consequences of the courses of action that could be taken.
- *Practice* – anyone with managerial experience, or who is approaching careers choices, understands the importance of developing practical knowledge and skills that can be applied in practice. This requires an approach, as well as frameworks and techniques, that can be adapted to take account of the complexity and ambiguity of operations, yet give guidance to identifying and implementing potential solutions.

Who should use this text?

This text is intended to provide an introduction to operations and process management for everyone who wishes to understand the nature, principles and practice of the subject. It is aimed primarily at those who have some management experience (although no prior academic knowledge of the area is assumed), or who are about to embark on a career in management. For example:

- *MBA students* should find that its practical discussions of operations management activities enhance their own experience.
- *Postgraduate students* on other specialist masters degrees should find that it provides them with a well-grounded and, at times, critical approach to the subject.
- *Executives* should find its diagnostic structure helpful to provide an understandable route through the subject.

What makes this text distinctive?

It has a clear structure

The text is structured on a model of operations management that distinguishes between activities that contribute to the direction, design, delivery and development of operations and processes.

It is based on practical diagnostic logic

Every chapter follows a series of questions that forms a 'diagnostic logic' for the topic. These are the questions that anyone can ask to reveal the underlying state of their, or any other, operations. The questions provide an aid to diagnosing where and how an operation can be improved.

It is illustrations-based

Operations management is a practical subject and cannot be taught satisfactorily in a purely theoretical manner. Because of this, each chapter starts with a real-life example of how the topic is treated in practice and provides additional examples in relation to specific issues within each chapter.

It identifies key operations principles

Whenever a core idea of operations and process management is described in the text, a brief 'operations principle' summary is included in the margin. This helps to distil those essential points of the topic.

It includes critical commentaries

Not everyone agrees about what is the best approach to the various topics and issues within the subject. This is why we have, at the end of each chapter, included a 'critical commentary'. These are alternative views to the one being expressed in the main flow of the text. They do not necessarily represent our view, but they are worth debating.

Each chapter includes summary checklists

Each chapter is summarised in the form of a list of checklist questions. These cover the essential questions that anyone should ask if they wish to understand the way their own or any other operation works. More importantly, they can also act as prompts for operations and process improvement.

Each chapter finishes with a case study

Every chapter includes a case study, relating real or realistic situations that require analysis, decision, or both. The cases have sufficient content to serve as the basis of case sessions in class, but are short enough to serve as illustrations for the less formal reader.

Each chapter includes an 'applying the principles' section

Selected problems, short exercises and activities are included at the end of each chapter. These provide an opportunity to test out your understanding of the principles covered in the chapter.

Each chapter includes a 'taking it further' section

A short annotated list of further reading and useful websites is provided, which take the topics in the chapter further, or treat some important related issues.

Suggested 'model answers' are available for all the 'applying the principles' exercises

Answers to the first two questions are available on the companion website for students. Answers to all the questions are available to bone fide tutors and lecturers.

Instructor's manual and PowerPoint slides

Visit go.pearson.com/uk/he/resources to find valuable online resources. A dedicated updated web-based instructor's manual is available to lecturers adopting this text. It includes teaching notes for all chapters, guided solutions for all case studies in the text, guided solutions for active cases and ideas for teaching them. A set of PowerPoint slides featuring figures and illustrations from the main text is also available.

About the authors

NIGEL SLACK is an Emeritus Professor of Operations Management and Strategy at the University of Warwick, an Honorary Professor at the University of Bath and an Honorary fellow of the European Operations Management Association. Previously he has been Professor of Service Engineering at the University of Cambridge, Professor of Manufacturing Strategy at Brunel University London, a University Lecturer in Management Studies at the University of Oxford and Fellow in Operations Management at Templeton College, Oxford. He worked initially as an industrial apprentice in the hand-tool industry and then as a production engineer and production manager in light engineering. He holds a Bachelor's degree in Engineering and Master's and Doctor's degrees in Management, and is a Chartered Engineer. He is the author of many books in Operations Management, including *The Operations Advantage*, published in 2017, *Operations Management* (with Alistair Brandon-Jones), the ninth edition published in 2019, *Operations Strategy* (with Michael Lewis), the sixth edition published in 2020, *The Manufacturing Advantage*, published in 1991, *Making Management Decisions* (with Steve Cooke), published in 1991, *Service Superiority* (with Robert Johnston), published in 1993, *The Blackwell Encyclopedic Dictionary of Operations Management*, published in 1997 and *Perspectives in Operations Management* (with Michael Lewis), published in 2003. Nigel has authored numerous academic papers and chapters in books. He also acts as a consultant to many international companies around the world in many sectors, especially financial services, transport, leisure and manufacturing. His research is in the operations and manufacturing flexibility and operations strategy areas.

ALISTAIR BRANDON-JONES is a Full Chaired Professor in Operations and Supply Chain Management at the University of Bath, School of Management. He is also an Adjunct Professor for Hult International Business School and Danish Technical University. Since re-joining Bath in 2014, he has served as Associate Dean for Post-Experience Education, responsible for MBA, EMBA and DBA programmes, and is currently Head of the Information, Decisions, and Operations Division (IDO). Previously, he was a Reader at Manchester Business School, Associate Professor and Assistant Professor at the University of Bath and a Teaching Fellow at Warwick Business School, where he also completed his PhD. His other books include *Operations Management*, the ninth edition published in 2019, *Essentials of Operations Management*, the second edition published in 2018 and *Quantitative Analysis in Operations Management*, published in 2008. Alistair is an active empirical researcher, focusing on digitisation in operations and supply networks, inter-organisational relationships and healthcare operations. His research has been published extensively in world-elite journals including *Journal of Operations Management, International Journal of Operations and Production Management, International Journal of Production Economics* and *International Journal of Production Research.* Alistair has led Operations Management, Operations Strategy, Supply Chain Management and Project Management courses at all levels and has been invited to teach at various international institutions, including the University of Cambridge, SDA Bocconi, Warwick Business School, NOVA University, Danish Technical University, Hult International Business School, Warwick Medical School and University College Dublin. His teaching has been recognised with a number of prizes for excellence and contributions to pedagogy, including from the Times Higher Education, Association of MBAs (AMBA), Production Operations Management Society (POMS), University of Bath, University of Manchester, University of Warwick and Hult International Business School. Alistair has consulting and executive development experience with organisations in various sectors, including petrochemicals, health, financial services, manufacturing, defence, government and non-profit. He is also a non-executive director for Brevio, a technology platform start-up aimed at improving grant processes for charities and grant makers.

Author's acknowledgements

In preparing this text, the authors, as usual, unashamedly exploited their friends and colleagues. In particular we had invaluable help from a great and distinguished reviewer team and from colleagues who have provided valuable feedback on various aspects of the project. For their help in case writing, we particularly thank: Vaggelis Giannikas, University of Bath, UK; Jas Kalra, University College London, UK; Jens Roehrich, University of Bath, UK; Nigel Spinks, Henley Business School, UK; and Brian Squire, University of Bath, UK.

Colleagues at many universities have helped and contributed to our understanding of the subject, including: Malcolm Afferson, Sheffield Hallam University, UK; Pär Åhlström, Stockholm School of Economics, Sweden; David Bamford, University of Huddersfield, UK; Vikram Bhakoo, University of Melbourne, Australia; Mattia Bianchi, Stockholm School of Economics, Sweden; Umit Bitici, Strathclyde University, UK; Briony Boydell, Portsmouth University, UK; Raffaella Cagliano, Politecnico di Milano, Italy; Dan Chicksand, Aston University, UK; Ben Clegg, Aston University, UK; Paul Coughlan, Trinity College Dublin, Ireland; Pamela Danese, University of Padova, Italy; Marc Day, Reading University, UK; Stephen Disney, Exeter University, UK; Des Doran, Kingston University, UK; Joy Field, Boton College, USA; Paul Forrester, Keele University, UK; Gino Franco, Derby University, UK; Abby Ghobadian, Reading University, UK; Janet Godsell, Warwick University, UK; Roger Hall, Huddersfield University, UK; Ingjaldur Hannibalsson, University of Iceland, Iceland; Linda Hendry, Lancaster University, UK; Matthias Holweg, Oxford University, UK; Koos Krabbendam, University of Twente, The Netherlands; Bart MacCarthy, Nottingham University, UK; John Maguire, Sunderland University, UK; Ian McCarthy, Simon Fraser University, Canada; Ronnie Mcmillan, University of Strathclyde, UK; Harvey Maylor, Cranfield University, UK; Phil Morgan, Oxford Brookes University, UK; Andy Neely, Cambridge University, UK; Steve New, Oxford University, UK; Jan Olhager, Lund University, Sweden; Carie Queenan, University of South Carolina, USA; Zoe Radnor, City, University of London, UK; Pietro Romano, Udine University, Italy; Sarah Schiffling, Lincoln University, UK; Mike Shulver, Bath University, UK; Alison Smart, University of Glasgow, UK; Rui Sousa, Universidade Católica Portuguesa, Portugal; Martin Spring, Lancaster University, UK; Venu Venugopal, Nyenrode University, The Netherlands; Ann Vereecke, Vlerick Business School, Belgium; Rohit Verma, Cornell Univerity, USA; Jan de Vries, University of Groningen, The Netherlands; Graham Walder, Wolverhampton University, UK; Helen Walker, Cardiff University, UK; and Maggie Zeng, Kent University, UK.

Our very particular thanks go to our academic colleagues in the Operations Management Group at Warwick Business School, and Information, Decisions and Operations Division in Bath School of Management. They helped, both by contributing ideas and by creating a lively and stimulating work environment.

Our thanks to colleagues at Warwick go to: Vikki Abusidualghoul; Haley Beer; Nicola Burgess; Mehmet Chakkol; Max Finne; Emily Jamieson; Mark Johnson; Pietro Micheli; Giovanni Radaelli; Ross Ritchie; Rhian Silvestro; and Chris Voss.

Our thanks to colleagues at Bath go to: Meriem Bouazzaoui; Teslim Bukoye; Melih Celik; Soheil Davari; Britt Davidson; David Ellis; Günes Erdogan; Emmanuel Fragniere; Vaggelis Giannikas; Elvan Goklap; Andrew Graves; Joanne Hinds; Adam Joinson; Richard Kamm; Gilbert Laporte; Michael Lewis; Sheik Meeran; Meng Meng; Fotios Petropoulos; Lukasz Piwek; Jens Roehrich; Ece Sanci; Gamila Shoib; Michael Shulver; Brian Squire; Oliver Thomasson; Christos Vasilakis; and Baris Yalabik.

Many practitioners also contributed, not least by keeping us honest to what really works in real operations. We thank and acknowledge: Vian Davys, CLP Holdings; David Garman, Deputy Chairman, John Menzies plc; Paul Gaskell, MD, The Value Engineers; and Tony Harpur, Ramblers Walking Holidays.

The prerequisite for any text of this type is that it serves a real market need. We were privileged to receive advice from the Pearson team; some of the most insightful people in business education publishing, in particular, Anita Atkinson, Felicity Baines and Rufus Curnow. Finally, to our families, Angela and Kathy, and Emma, Noah and George, your ongoing support is invaluable.

Nigel Slack and Alistair Brandon-Jones

1 Operations and processes

Introduction

Operations management is about how organisations produce goods and services. Every product you use and every service you experience comes to you courtesy of the operations managers who organised its production. They may not always be called operations managers, but that is what they really are. They are the people who design, run and improve the processes that produce services and products for their customers. But operations *and process* management is even wider than this. After all, managers in other functions, such as marketing, sales and finance, also manage processes. They supply internal 'customers' with services such as marketing plans, sales forecasts, budgets, and so on. In fact, all parts of all organisations are made up of processes. That is why operations and process management is of direct relevance to all managers, irrespective of what type of organisation they work for, or which function they work in. And that is what this text is about – the tasks, issues and decisions that are necessary to manage processes effectively, both within the operations function and in other parts of the business where effective process management is equally important. This first chapter is an introductory chapter, so we will examine some of the basic principles of operations and process management. The model that is developed to explain the subject is shown in Figure 1.1.

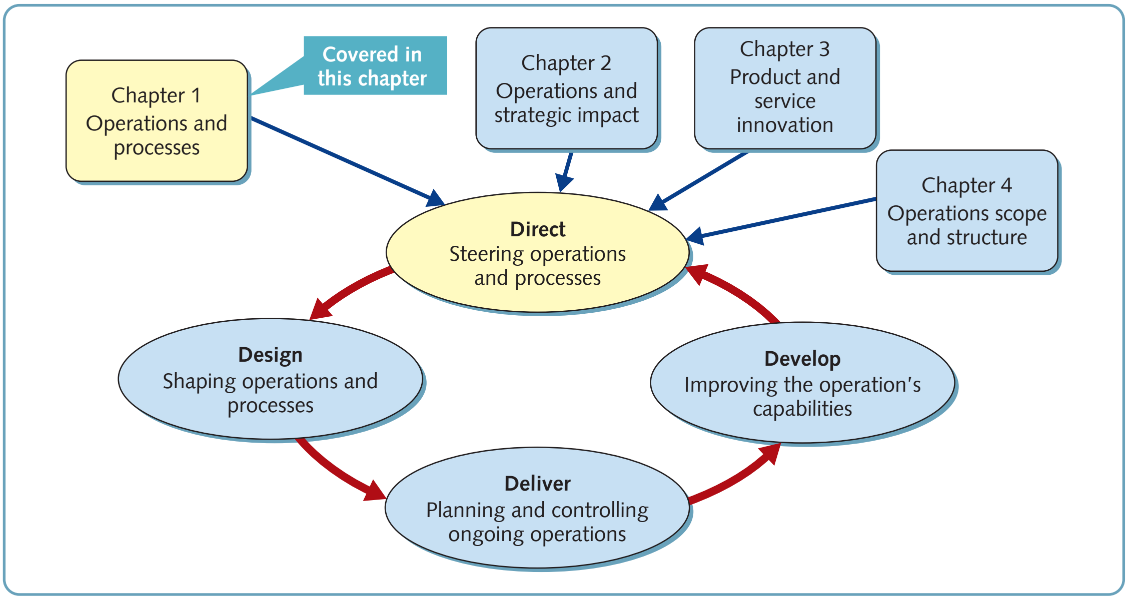


Figure 1.1 **Operations and process management is about how organisations produce goods and services**

EXECUTIVE SUMMARY

1.1 Does the organisation understand the potential of operations and process management?

↓

1.2 Does the organisation take a process perspective?

↓

1.3 Does operations and process management have a strategic impact?

↓

1.4 Are processes managed to reflect their operating circumstances?

↓

1.5 Is operations and process decision-making appropriate?

Each chapter has a short executive summary structured around the number of diagnostic questions used in the chapter. These diagnostic questions represent the basic line of enquiry that can reveal the nature and relevance of the topic covered in each chapter.

1.1 Does the organisation understand the potential of operations and process management?

The operations function is the part of the organisation that produces products or services. Every organisation has an operations function because every organisation produces some mixture of products and services. It is a central and important activity for *any* organisation. 'Operations' is not always called by that name, but whatever its name, it is always concerned with managing the core purpose of the business – serving its customers by producing some mix of products and services. 'Processes' also produce products and services, but on a smaller scale. They are the component parts of operations. But other functions also have processes that need managing. In fact *every* part of *any* business is concerned with managing processes. All managers have something to learn from studying operations and process management, because the subject encompasses the management of all types of operation, no matter in what sector or industry, and all processes, no matter in which function.

1.2 Does the organisation take a process perspective?

A 'process perspective' means understanding businesses in terms of all their individual processes. It is only one way of modelling organisations, but it is a particularly useful one. Operations and process management uses the process perspective to analyse businesses at three levels: the operations function of the business; the higher and more strategic level of the supply network; and at a lower, more operational level of individual processes. Within the business, processes are only what they are defined as being. The boundaries of each process can be drawn as thought appropriate. Sometimes this involves radically reshaping the way processes are organised, for example, to form end-to-end processes that fulfil customer needs.

1.3 Does operations and process management have a strategic impact?

Operations and process management can make or break a business. When they are well managed, operations and processes can contribute to the strategic impact of the business in four ways: cost, revenue, investment and capabilities. Because the operations function has responsibility for much of a business's cost base, its first imperative is to keep costs under control. Additionally, it should be looking to enhance the business's ability to generate revenue through the way it provides service and quality. Furthermore, all failures are ultimately process failures; well-designed processes should have less chance of failing and more chance of recovering quickly from failure. Also, because operations are often the source of much investment, it should be aiming to get the best possible return on that investment. Finally, the operations function should be laying down the capabilities that will form the long-term basis for future competitiveness.

1.4 Are processes managed to reflect their operating circumstances?

Not necessarily. Processes differ, particularly in what are known as the four Vs: volume, variety, variation and visibility. High-volume processes can exploit economies of scale and be systematised. High-variety processes require enough inbuilt flexibility to cope with the wide variety of activities expected of them. High-variation processes must be able to change their output levels to cope with highly variable and/or unpredictable levels of demand. High-visibility processes add value while the customer is 'present' in some way and therefore must be able to manage customers' perceptions of their activities. Generally, high volume together with low variety, variation and visibility facilitate low-cost processes, while low volume together with high levels of variety, variation and visibility all increase process costs. Yet in spite of these differences, operations managers use a common set of decisions and activities to manage them. These activities can be clustered under four groupings: directing the overall strategy of the operation; designing the operation's products, services and processes; planning and controlling process delivery; and developing process performance.

1.5 Is operations and process decision-making appropriate?

The range of operations decisions is wide and covers four broad areas that we categorise as: '*directing* the overall strategy of the operation'; '*designing* the operation's processes'; 'planning and control process *delivery*'; and '*developing* process performance'. However, there are always overlaps and interrelationships between the categories. Yet, no matter what type of decision, operations managers use models (many of which are included in this text) to help them make decisions. Some models are quantitative, some are qualitative, but in practice a blend of qualitative and quantitative is often the most useful approach. Remember, though, that all models are simplifications of a far more complex reality. Which is one reason for the interest in 'behavioural operations management', which attempts to incorporate real (usually non-rational) behaviour into operations decision-making.

1.1 Diagnostic question: *Does the organisation understand the potential of operations and process management?*

Operations and process management is the activity of managing the resources and processes that produce products and services, for internal and external customers. It is a central and important activity for *any* organisation. The core body of knowledge for the subject comes from 'operations management', which examines how the 'operations function' of an organisation produces products and services for external customers. In some organisations an operations manager could be called by some other name, for example, a 'fleet manager' in a logistics company, an 'administrative manager' in a hospital or a 'store manager' in a supermarket. Note also that throughout this text the terms 'the operation', 'the operations function', and 'operations', will be used, more or less interchangeably. Also 'the organisation', 'the business', 'the firm' and 'the enterprise' are used to mean whatever formal body (public or private) one is working for.

All enterprises have 'operations'

All organisations have 'operations', because all organisations produce products, services or some mixture of both. If you think that you don't have an operations function, you are wrong. If you think that your operations function is not important, you are also wrong. In most enterprises the operations function represents the bulk of its assets and the majority of its people. It is the means by which they serve their customers and provide an economic and/or social return for their stakeholders. An effective operations function has the potential to survive in a turbulent environment and the ability to maintain a steady improvement in its performance. By contrast, a poorly managed operations function, especially if it fails to provide adequate service to its customers or fails to provide the efficiency to work within its cost constraints, will always prevent an organisation from achieving its objectives, whether social or economic. But the subject does have something of an image problem. It is sometimes seen as dealing with routine, 'technical', low-level activities that obviously have to be done – but preferably by someone else. Worthy maybe, and even challenging, but neither exciting nor of direct interest to anyone outside the operations function itself. Wrong. 'Operations' is how one makes things happen. It is how organisations release whatever expertise they have, in order to create value. It is through one's operations that customers are served. It is through operations that one uses resources to their best advantage. And it is through one's operations that strategy is made into reality.

> **OPERATIONS PRINCIPLE**
> ***All organisations have 'operations' that produce some mix of products and services.***

But not all operations are the same

Look at the six businesses illustrated in Figure 1.2. There are two financial service companies, two manufacturing companies and two hotels. All of them have *operations functions* that produce the things that their customers are willing to pay for. Hotels produce accommodation services; financial services invest, store, move or sell money and investment opportunities; and manufacturing businesses physically change the shape and the nature of materials to produce products. These businesses are from different sectors (banking, hospitality and manufacturing), but it is not that they operate in different sectors of the economy that makes these businesses different from each other. There are often bigger differences *within* economic sectors than *between* them. The main difference between how their operations activities need to be managed is more closely related to the market position that they occupy. So, for example, all three operations in the left-hand column provide value-for-money products and services and compete largely on cost. The three in the right-hand column provide more 'upmarket' products and services that are more expensive to produce and compete on some combination

> **OPERATIONS PRINCIPLE**
> ***The economic sector of an operation is less important in determining how it should be managed than its intrinsic characteristics.***

Financial services

Tetra Images/Brand X Pictures/Getty Images

An account management centre at a large retail bank

Financial analyst advising a client at an investment bank

FangXiaNuo/E+/Getty Images

Furniture manufacturing

Albert Karimov/Alamy Stock Photo

Mass production of kitchen units

Craft production of reproduction 'antique' furniture

Alistair Berg/DigitalVision/Getty Images

Hotels

Chubykin Arkady/Shutterstock

Value-for-money hotel

Lobby of an international luxury hotel

Mint Images RF/Getty Images

Figure 1.2 All types of business have 'operations' because all businesses produce some mix of products and services. And the differences in the operations *within* a category of business are often greater than the differences *between* business sectors

of high specification and customisation. The implication of this is important. It means that the surface appearance of a business and its economic sector are less important to the way its operations should be managed than the intrinsic characteristics of what it is trying to achieve, such as the volume of its output, the variety of the products and services it needs to produce and, above all, how it is trying to compete in its market.

From production to operations management

Figure 1.3 illustrates how the scope of this subject has expanded over time. Originally, operations management was almost exclusively associated with the manufacturing sector. It would have been called 'production' or 'manufacturing' management. Starting in the 1970s and 1980s the term *operations management* was used to reflect two trends. First, and most importantly, it was used to imply that many of the ideas, approaches and techniques traditionally used in the manufacturing sector could be equally applicable in the (much larger) service sector. The second use of the term was to expand the scope of 'production' in manufacturing companies to include 'non-core', but important, production-related processes such as purchasing, physical distribution, after-sales service, product development and so on.

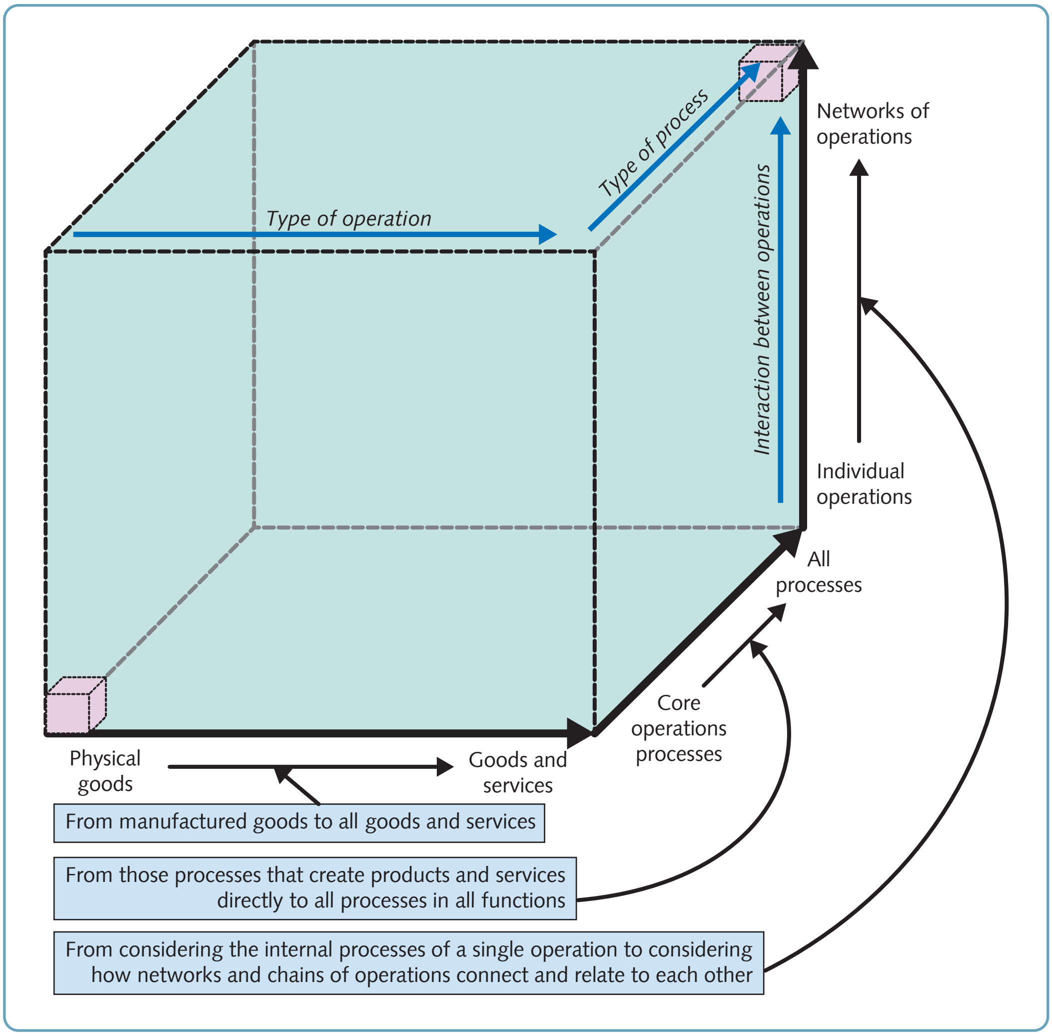


Figure 1.3 The scope of operations management has expanded from focusing on processes in manufacturing organisations to include service organisations, processes in other functions (such as marketing, finance and HRM) and how operations interact with each other in supply networks and chains

From operations management to operations and process management . . .

More recently the term *operations and process management* (or sometimes just process management) has been used to denote the shift in the scope of the subject to include the whole organisation. Within any business, the production of products and services is not confined to the operations function. After all, every part of any business achieves its objectives by organising its resources such as people, information systems, buildings and equipment into individual 'processes'. A 'process' is an arrangement of resources and activities that transform inputs into outputs that satisfy (internal or external) customer needs. For example, the marketing function 'produces' marketing plans and sales forecasts, the accounting function 'produces' budgets, the human resources function 'produces' development and recruitment plans, and so on. In fact *every* part of *any* business is concerned with managing processes.

The difference between *operations* and *processes* is one of scale, and therefore complexity. Both transform inputs into outputs (we shall look at this idea later), but processes are the smaller version. They are the component parts – the building blocks – of an operation. So, 'operations and process management' is the term we use to encompass the management of all

OPERATIONS PRINCIPLE
All operations are composed of processes. A process is an arrangement of resources and activities that transform inputs into outputs that satisfy (internal or external) customer needs.

types of operation, no matter in what sector or industry, and all processes, no matter in which function of the business. The general truth is that processes are everywhere, and all types of managers have something to learn from studying operations and process management. This is very much how we treat the subject in this text. That is why it is called *Operations and Process Management*: it includes an examination of the operations function in both manufacturing and service sectors, as well as the management of processes in non-operations functions.

. . . and operations and supply management

The third big development in how the subject was viewed was the (almost) universal acceptance that operations outside the conventional boundary of an enterprise would have an important effect on how it performed. The suppliers of goods and services to an operation could have as much impact on its activities as how it managed its internal resources and processes. Moreover, the boundary of an enterprise is not fixed or immutable, it is determined by decisions made by the operation itself. For example, should an operation perform activities itself, or should it outsource the activity? No operation does everything itself; some degree of outsourcing is inevitable. Therefore, confining this subject to internal activities would clearly give only a partial picture. This is why we frequently refer to operations and process-based issues in the wider 'supply network' (we shall define this term later) and devote two chapters specifically to supply network issues (Chapters 4 and 7).

Case example

Torchbox: award-winning web designers[1]

There are around two billion websites in the world. So many that we usually take them for granted, yet browsing websites as part of our studies, our job or our leisure is an activity that we all do; probably every day, probably many times each day. It's important. And, not surprisingly, there is a whole industry devoted to designing websites so that they have the right type of impact. In fact, taken over the years, web development has been one of the fastest-growing industries in the world. But it's also a tough industry. Not every web design company thrives, or even survives beyond a couple of years. To succeed, web designers need technology skills, design capabilities, business awareness and operational professionalism. One that has succeeded is Torchbox, an independent web design and development company, with offices in Oxfordshire, Bristol and Cambridge in the UK. Founded back in 2000 by Tom Dyson and Olly Willans, on Employee Ownership Day 2019 they handed over the ownership of the company to its employees, who elected a board of employee trustees to oversee the company's work and who have ultimate control over the business.

It was a move that fitted their ethos of providing high-quality, cost-effective and ethical solutions for clients who come primarily, but not exclusively, from the charity, non-governmental organisations and public sectors. *'There are a number of advantages about being a relatively small operation'*, says co-founder and Technical Director Tom Dyson, who has been responsible for the technical direction of all major developments. *'We can be hugely flexible and agile, in what is still a dynamic market. But at the same time we have the resources and skills to provide a creative and professional service. Any senior manager in a firm of our size cannot afford to be too specialised. All of us here have their own specific responsibilities; however, every one of us shares the overall responsibility for the firm's general development. We can also be clear and focused on what type of work we want to do. Our ethos is important to us. We set out to work with clients who share our commitment to environmental sustainability and responsible, ethical business practice; we take our work, and that of our clients, seriously.'*

Nevertheless, straightforward operational effectiveness is essential to Torchbox's business. *'We know how to make sure that our projects run not only on time and to budget'*, says Olly Willans, the firm's Creative Director, *'but we also like to think that we provide an enjoyable and stimulating experience – both for our customers' development teams and for our staff too. High standards of product and* *service are important to us: our clients want accessibility, usability, performance and security embedded in their web designs, and of course, they want things delivered on-time and on-budget. We are in a creative industry that depends on fast-moving technologies, but that doesn't mean that we can't also be efficient. We back everything we do with a robust feature-driven development process using a kanban project management methodology which helps us manage our obligations to our clients.'* Tom Dyson adds, *'Using sound operations management techniques helps us constantly to deliver value to our clients. We like to think that our measured and controlled approach to handling and controlling work helps ensure that every hour we work produces an hour's worth of value for our clients and for us.'*

Maskot/Getty Images

1.2 Diagnostic question: *Does the organisation take a process perspective?*

Central to making operations and process management a significant contributor to an enterprise's success is the idea of a 'process perspective'. A process perspective means understanding that all parts of the business can be seen as processes, and that all processes can be managed using operations management principles. Yet, although important, a process perspective is not the only way of describing businesses, or any type of organisation. One could represent an organisation as a conventional 'organisational structure' that shows the reporting relationships between various departments or groups of resources. But even a little experience in any organisation shows that rarely, if ever, does this fully represent the way the organisation actually works. Alternatively, one could describe an organisation through the way it makes decisions: how it balances conflicting criteria, weighs up risks, decides on actions and learns from its mistakes. Or, one could describe an organisation by explaining its culture – its shared values, ideology, pattern of thinking and day-to-day rituals, or its power relationships – how it is governed, seeks consensus (or at least reconciliation), and so on. Or, and this is the significant point, one can represent the organisation as a collection of processes, interconnecting and (hopefully) all contributing to fulfilling its strategic aims. This is the perspective that we emphasise throughout this text. As we define it here, the process perspective analyses businesses as a collection of interrelated processes. Some of these processes will be within the operations function, and will contribute directly to the production of its products and services. Other processes will be in the other functions of the business, but will still need managing using similar principles to those within the operations function.

None of these various perspectives on organisations gives a total picture. Each perspective adds something to our ability to understand, and therefore more effectively manage a business. Nor are these perspectives mutually exclusive. A process perspective does not preclude understanding the influence of power relationships on how processes work, and so on. We use the process perspective here, not because it is the *only* useful and informative way of understanding businesses, but because it is the perspective that directly links the way we manage resources in a business with its strategic impact. Without effective process management, the best strategic plan can never become reality. The most appealing promises made to clients or customers will never be fulfilled. In addition, the process perspective has traditionally been undervalued. The subject of operations and process management has only recently come to be seen as universally applicable and, more importantly, universally valuable.

OPERATIONS PRINCIPLE
There are many valid approaches to describing organisations. The process perspective is a particularly valuable one.

Operations and process management is relevant to all parts of the business

If processes exist everywhere in the organisation, operations and process management will be a common responsibility of all managers, irrespective of which function they are in. Each function will have its 'technical' knowledge, of course. In marketing, this includes the market expertise needed for designing and shaping marketing plans; in finance, it includes the technical knowledge of financial reporting conventions. Yet each will also have an *operations* role that entails using its processes to produce plans, policies, reports and services. For example, the marketing function has processes with inputs of market information, staff, computers and so on. Its staff transforms the information into outputs such as marketing plans, advertising campaigns and sales force organisation. In this sense, all functions are operations with their own collection of processes. The implications of this are very important. As every manager, in all parts of an organisation is, to some extent, an operations manager, they all should want to give good service to their (internal) customers, and they all should want to do this efficiently. So, operations management must be relevant for all functions, units and groups within the organisation. And the concepts, approaches and techniques of operations management can be used to help improve any process in any part of the organisation.

> OPERATIONS PRINCIPLE
> ***All parts of the business manage processes so all parts of the business have an operations role and need to understand operations management.***

The 'input–transformation–output' model

> OPERATIONS PRINCIPLE
> ***All processes have inputs of transforming and transformed resources that they use to create products and services.***

Central to understanding the processes perspective is the idea that all processes and operations transform *inputs* into *outputs*. Figure 1.4 shows the *general transformation process model* that is used to describe the nature of processes and operations. Put simply, processes and operations take in a set of input resources, some of which are transformed into outputs of products and/or services and some of which do the transforming.

Input resources

Transformed resource inputs are the resources that are changed in some way within a process. They are usually materials, information or customers. For example, one process in a bank prints statements of accounts for its customers. In doing so, it is processing materials. In the bank's branches, customers are processed by giving them advice regarding their financial affairs, making payments, etc. However, behind the scenes, most of the bank's processes are concerned

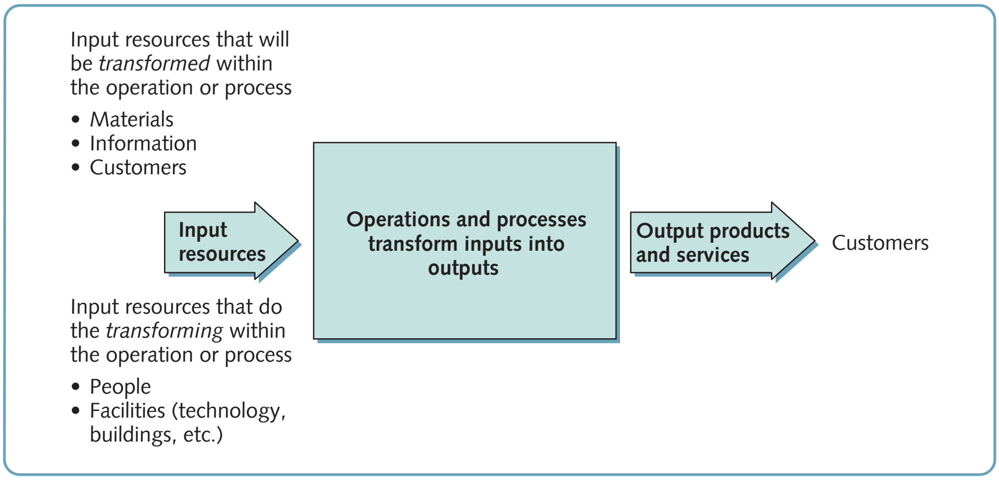


Figure 1.4 **All processes and operations are input–transformation–output systems that use 'transforming' resources to work on 'transformed' resources in order to produce products and services**

> **OPERATIONS PRINCIPLE**
> ***Transformed resource inputs to a process are materials, information or customers.***

> **OPERATIONS PRINCIPLE**
> ***All processes have transforming resources of facilities (equipment, technology, etc.) and people.***

with processing information about its customers' financial affairs. In fact, for the bank's operations function as a whole, it is its information transforming processes that are probably the most important. As customers, we may be unhappy with a clunky website or if we are not treated appropriately in the bank branch. But if the bank makes errors in our financial transactions, we suffer in a far more fundamental way.

There are two types of *transforming* resource that form the 'building blocks' of all processes. They are *facilities* – the buildings, equipment, plant and process technology of the operation, and *people* – who operate, maintain, plan and manage the operation.

The exact nature of both facilities and people will differ between processes. In a five-star hotel, facilities consist mainly of buildings, furniture and fittings. In a nuclear-powered aircraft carrier, its facilities are the nuclear generator, turbines and sophisticated electronic equipment. Although one operation is relatively 'low-technology' and the other 'high-technology', their processes all require effective, well-maintained facilities. Staff will also differ between processes. Most staff employed in a food processing plant may not need a very high level of technical skill, whereas most staff employed by an accounting firm in an audit process are highly skilled in their own particular 'technical' skill (accounting). Yet although skills vary, all staff have a contribution to make to the effectiveness of their operation. A food factory worker who does not follow hygiene procedures will damage the business and its customers just as surely as an accountant who cannot add up.

Front- and back-office transformation

A distinction that is worth noting at this point, mainly because it has such an impact on how operations and processes are managed, is that between 'front-' and 'back-office' transformation. The 'front-office' (or 'front-of-house') parts of an operation are those processes that interact with (transform) customers. 'Back-office' (or 'back-of-house) operations are the processes that have little or no direct contact with customers, but perform the activities that support the front office in some way. The distinction is illustrated in Figure 1.5. But, as implied by the figure, the boundary between front and back offices is not clean. Different processes within an operation could have different degrees of exposure (what we refer to later as 'visibility') to customers.

Output for customers

All processes and operations produce products and/or services, and although products and services are different, the distinction can be subtle. We will look at the differences between products and services later in this chapter. But whether classed as a product or service, the outputs from processes and operations are (or should be) intended to serve the requirements of customers. For the operations function of a business as a whole, the customers who need to be served are likely to be *external* customers – those parties or organisations that use the

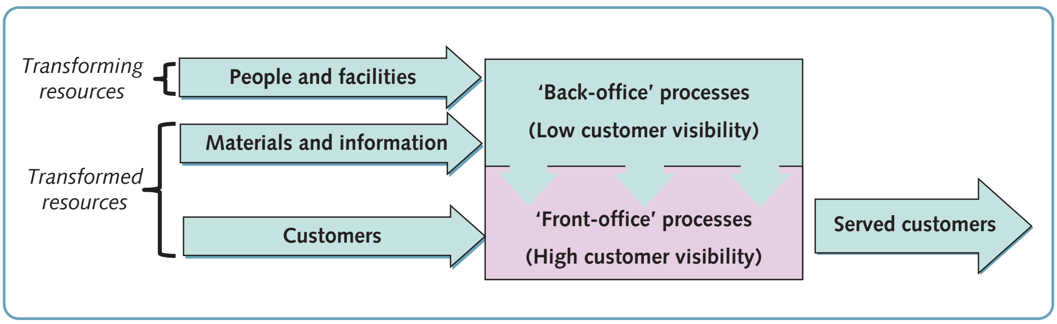


Figure 1.5 **When the main transformed resource is the customers themselves, it is useful to distinguish between 'front-office' processes that act on customers directly and 'back-office' processes that provide indirect services**

Case example

Marina Bay Sands[2]

There are very few better examples of how back and front offices work together than the hotel industry. As customers, we naturally judge a hotel primarily on its front-office, client-facing staff and facilities. But without effective back-office operations, customers would soon find that their front-office experience would be very much affected. This is certainly true for the Marina Bay Sands hotel in Singapore. Located in the heart of Singapore's Central Business District, Marina Bay Sands is an integrated, multi-award-winning, luxury resort owned by the Las Vegas Sands corporation, incorporating a hotel with over 2,500 rooms, a huge convention and exhibition centre, restaurants, a shopping mall, museum, two large theatres and the world's largest atrium casino. The hotel's three towers are crowned by the spectacular Sands SkyPark, which offers a 360-degree view of Singapore's skyline. It is home to restaurants, lush gardens and the world's longest public cantilever housing an observation deck. Everything that guests experience is luxurious, from the sculptures by world-renowned artists to the infinity pool, the length of three Olympic swimming pools, mounted in the SkyPark on the very top of the hotel.

But the meticulous service provided by the hotel's highly trained front-of-house staff could not happen without the many back-of-house processes that customers do not always notice. Some of these processes are literally invisible to customers, for example those that keep the accounts, or those that maintain the hotel's air-conditioning systems, or the dim sum preparation (dim sum are steamed dumplings served in small, bite-sized portions). Yet these processes are all important – and mass operations in their own right. The hotel's specialist chefs prepare 5,000 individual pieces of its famous dim sum by hand every day. Some back-of-house departments rely more on technology. The hotel's laundry must clean and press tens of thousands of pieces every day. Four thousand pool towels alone need cleaning every day, as do 35 items of room linen. Which is a problem for an organisation whose sustainability policy commits it to minimising its use of water. It took an investment of over £10 million in water-saving technology to reduce the hotel's usage by 70 per cent. Other back-of-house operations have a direct impact on how customers view the hotel. For example, the wardrobe department that keeps the hotel's over-9,000 staff looking smart is reputed to be the most high-tech in the world. Its 18 automated conveyors each have slots for 620 individual items of uniform, all of which have individual identification chips so that they can be tracked. When someone wants to collect a freshly laundered uniform, they punch their staff number into a keypad and, behind the scenes, the conveyor system automatically delivers the uniform to the waiting staff member.

Some processes straddle the front/back-of-house divide. The valet parking operation parks up to 200 cars each hour in its 2,500 parking spaces, and retrieves them in a target retrieval time of seven minutes. Housekeeping cleans, tidies and stocks all the bedrooms. The hotel's 50 butlers serve the more exclusive suites and cater for a wide variety of demands (one guest asked them to arrange a wedding banquet at four hours' notice). It is a role that demands dedication and attention to detail. It is they who serve the in-room food prepared in the kitchens, and arrange the flowers cut and selected by the hotel's florists.

Vichy Deal/Shutterstock

business's products or services. However, for the individual processes within an operation, customers are usually (although not always) *internal* customers. Internal customers can be other processes, individuals or groups within the business; in fact anyone who is affected by the product or service.

The obvious difference between serving external and internal customers is their freedom of choice. External customers provide an organisation directly with its reason for being. They are the ones who provide the revenue (directly or indirectly) on which the organisation depends. Consistently failing to satisfy external customers is likely to result in them going elsewhere.

There is an obvious and clear reason for any operation to pay serious attention to how it serves external customers. Internal customers, at least in the short term, do not usually have the option of 'going elsewhere'. As the customer of your organisation's payroll process, you cannot choose to be served by an alternative payroll process if they make mistakes on your salary payments. Yet, although it may seem to have a less immediate impact, serving internal customers is, in the longer term, just as important. Poor service to internal customers will eventually result in poor service to external customers, or excessive costs to the organisation, or both.

SIPOC analysis

Do not dismiss the idea of the 'input–transformation–output' model as either too obvious or too trivial, it can be the basis of a useful first step in understanding and improving processes. This is sometimes called SIPOC analysis. SIPOC stands for suppliers, inputs, process, outputs and customers. As its name implies, it is a method of formalising what a process requires for its inputs, where those inputs come from, what the process entails (at a relatively general rather than a detailed level), what the process produces as outputs and for whom the outputs are intended. Figure 1.6 shows a simple example that describes a recruitment process performed by the human resources function of a company. The advantage of such an analysis is that it helps all those involved in the process to understand (and, more importantly, agree) what it involves and where it fits within the business. More than this, it can prompt important questions that can sometimes be overlooked. For example, exactly what information should suppliers to the process provide? What notice should they be required to give? In what form should the information be given? What are the important steps in the process and who is responsible for them? Who needs to be informed when the process is complete? How should its success be judged? And so on.

The network connection – three levels of analysis

The concept of the input–transformation–output process is, arguably, the first big idea that defines operations and process management. The second is that all operations processes are

S suppliers	I inputs	P process	O output	C customers
Functional head Project manager Recruitment agency	Recruitment request Long-term resourcing plan Labour market reports	Agree essential and desirable characteristics ↓ Authorise budget and timescale ↓ Formally contract recruitment agency ↓ Agree shortlist of candidates ↓ Make offer and negotiate conditions ↓ Confirm start date and arrange induction	Qualified and experienced new staff member Completed staff records Payment authorisation	Functional head Project manager New recruit

Figure 1.6 A simple SIPOC analysis for a recruitment process performed by the human resources function of a company

connected in networks. Earlier, we explained how the subject had developed to include consideration of whole supply networks. It is a powerful idea that emphasises the essential interconnectedness of operations and process management decisions. The network concept can also be applied at different levels of analysis – three in particular.

> OPERATIONS PRINCIPLE
> *A process perspective can be used at three levels: the level of the operation itself, the level of the supply network and the level of individual processes.*

The most obvious level is that of the business itself or, more specifically, the operations function of the business. (The other functions of the business could also be treated at this level.) And, while analysing the business at the level of 'the operation' is important, for a more comprehensive assessment we also need to analyse the contribution of operations and process management at a higher and more strategic level (the level of its whole supply network) and at a lower more operational level (the level of individual processes). These three levels of operations analysis are shown in Figure 1.7.

The process perspective at the level of the operation

The operations part of a business is itself an input–transformation–output system, which transforms various inputs to produce (usually) a range of different products and services. Table 1.1 shows some operations described in terms of their main resource inputs, the processes within the operations and the outputs from the operation. Note how some of the inputs to the operation are transformed in some way, while other inputs do the transforming. Note also how in some operations customers themselves are inputs. (The airline, department store and police department are all like this.) This illustrates an important distinction between operations whose

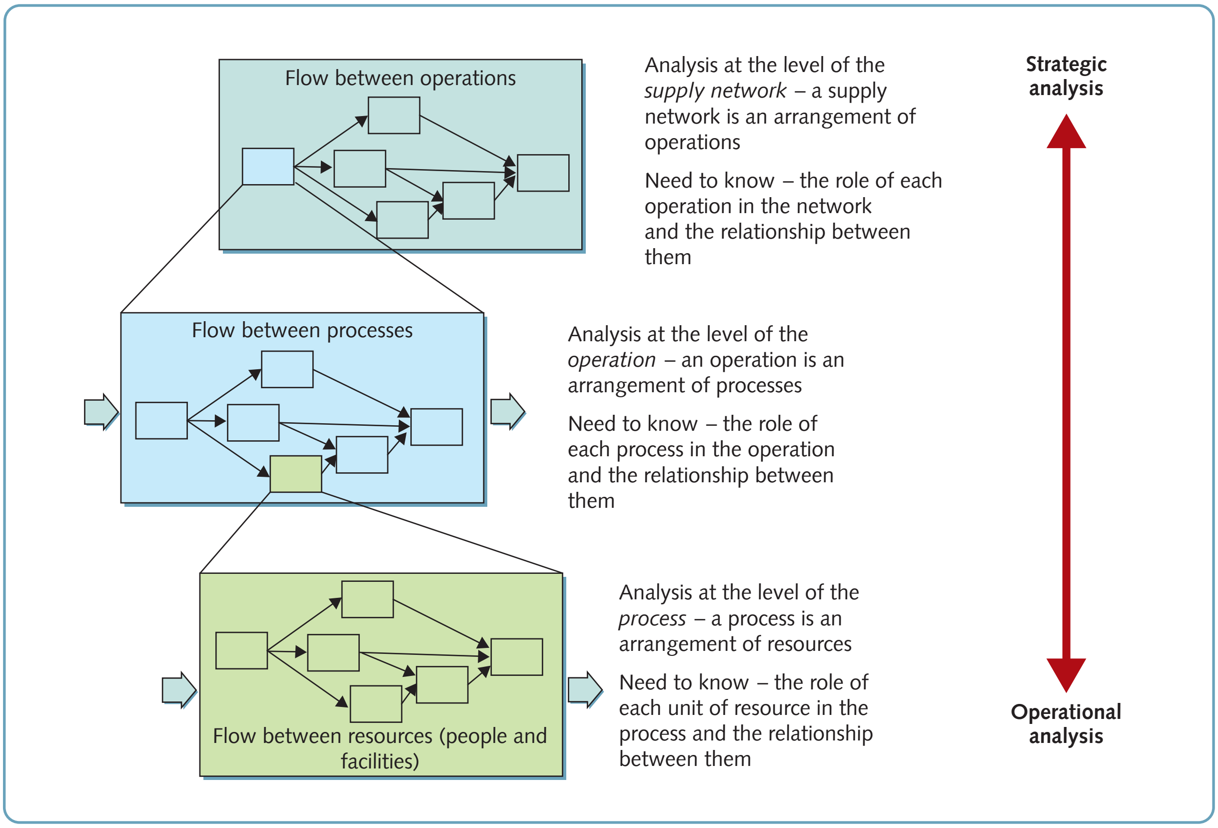


Figure 1.7 Operations and process management requires analysis at three levels – the supply network, the operation and the process

Table 1.1 Some operations described in terms of their inputs, purpose and outputs

Type of operation	*What are the operation's inputs?*	*What the processes within the operation?*	*What are the operation's outputs?*
Airline	Aircraft Pilots and air crew Ground crew *Passengers* *Cargo*	Passenger check-in assistance Baggage drop Security/seat check Board passengers Fly passengers and freight around the world Flight scheduling In-flight passenger care Transfer assistance Baggage reclaim, etc.	Transported passengers and freight
Department store	*Goods for sale* Staff sales Computerised registers *Customers*	Source merchandise Manage inventory Display products Give sales advice Sales Aftercare Complaint handling Delivery service, etc.	Customers and goods 'assembled' together
Police department	Police officers Computer systems *Information* *Public (law-abiding and criminal)*	Crime prevention Crime detection Information gathering/collating Victim support Formally charging/detaining suspects Managing custody suites Liaising with court/justice system, etc.	Lawful society Public with feeling of security
Frozen food manufacturer	*Fresh food* Operators Food-processing equipment Freezers	Source raw materials Input quality checks Prepare ingredients Assemble products Pack products Fast-freeze products Quality checks Finished goods inventory, etc.	Frozen food

Note: input resources that are transformed are shown in *italics*.

customers receive their outputs without seeing inside the operation, and those whose customers are inputs to the operation and therefore have some visibility of the operation's processes. Managing high-visibility operations where the customer is inside the operation usually involves a different set of requirements and skills to those whose customers never see inside the operation. (We will discuss the issue of visibility later in this chapter.)

Products or services

The difference between a 'product' and a 'service' is not always obvious and has provoked a lot of (not always useful) academic debate. At an obvious, but simple, level, a product is a physical and tangible thing (you can touch a car, a television or a phone). By contrast, a service is an activity that usually involves interaction with a customer (as with a physician and their patient) or something representing the customer (as with a package delivery service). The resources that carry out these services may be tangible, but not the service they provide. For many years the accepted distinction between products and services was not confined to intangibility, but included other characteristics of services abbreviated to 'IHIP', standing for:

- Intangibility, in that services are not physical items.
- Heterogeneity, in that they are difficult to standardise because each time a service is delivered, it will be different because the needs and behaviour of customers will, to some extent, vary.
- Inseparability, in that their production and consumption are simultaneous. The service provider (who 'produces' the service) is often physically present when its consumption by a customer takes place.
- Perishability, in that they cannot be stored because they have a very short 'shelf life'. They may even perish in the very instant of their creation, such as in the case of a theatre performance.

However, there are several problems with using these characteristics to define a 'service' – hence the academic debate. It is certainly not difficult to find examples of services that do not conform to them. Technology has also had a significant effect, both on the extent to which the IHIP characteristics apply and how the limits that they place on service operations can be overcome. In particular, the development of information and communication technology has opened up many new types of service offerings. Yet, although they cannot totally define what is a 'service' and what is a 'product', each of the IHIP characteristics does have some validity. Some operations do produce 'pure' products that exhibit few or none of these characteristics (for example, mineral extraction), while others produce 'pure' services that exhibit all of them (for example, a psychotherapy clinic).

Most operations produce outputs somewhere on a spectrum of the IHIP characteristics

Some operations produce just products and others just services, but most produce outputs that are a blend of the two. Figure 1.8 shows a number of operations positioned in a spectrum using the IHIP characteristics, from almost 'pure' goods producers to almost 'pure' service

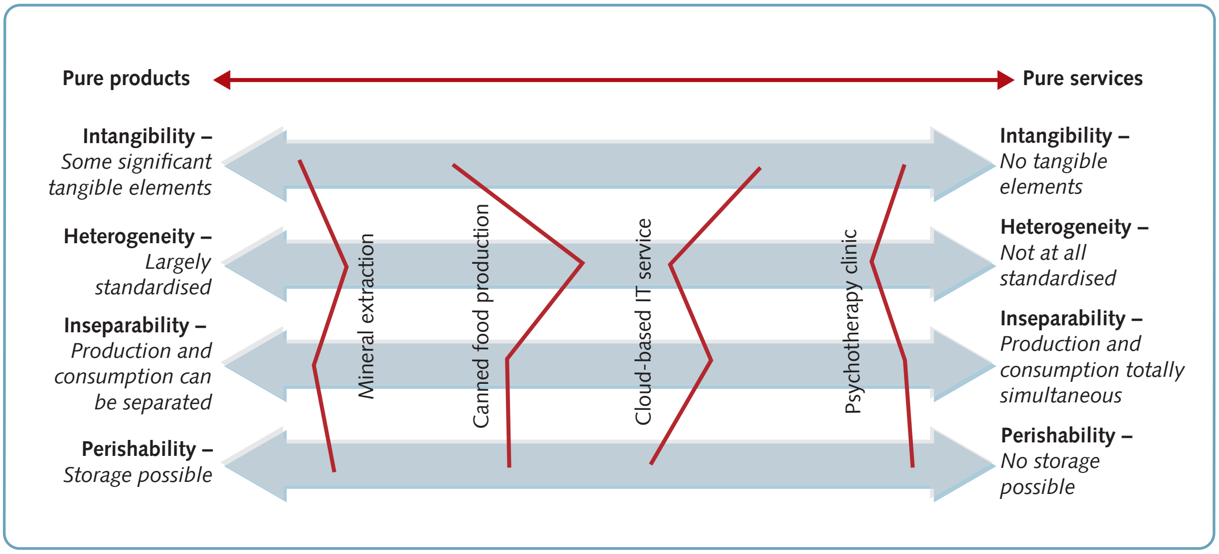


Figure 1.8 **Relatively few operations produce either purely products or purely services. The output from most types of operation blend the characteristics of 'pure' goods and 'pure' services**

> **OPERATIONS PRINCIPLE**
> ***Most operations produce a blend of tangible products and intangible services.***

producers. Mineral extraction operations (miners) are concerned almost exclusively with the product that comes from their mines. It is tangible, almost totally standardised, produced away from its consumption and storable. Canned food is similar – tangible, if not as standardised, but produced separately from its consumption and storable. However, customers (supermarkets) are also buying information concerning demand and delivery that makes the whole package less of a pure product. Closer to the pure service end of the spectrum, a cloud-based IT service has few, if any, tangible elements but may offer some customisation, and if customers want immediate access to the service, it may be able to store their preferences and data. Finally, some pure services conform almost totally to IHIP characteristics. A psychotherapy clinic, for example, provides personalised and close-contact therapeutic treatment for its customers without any tangible elements.

Using IHIP characteristics to distinguish between different types of output is of more than theoretical interest; they have real operational consequences. For example:

- Intangibility means it is difficult to define the 'boundary' of the less tangible elements of service. It therefore becomes particularly important to manage customers' expectations as to what the service comprises.
- Heterogeneity means that every service is different and difficult to standardise. Customers could ask for elements of service that are difficult to predict and may be outside the operation's capabilities. Cost efficiencies become difficult and staff must be trained to cope with a wide variety of requests.
- Inseparability means that production and consumption are simultaneous, so to meet all demand, operations must have sufficient capacity in place to meet demand as it occurs. However, customer guidance can reduce the need for contact (e.g. the use of FAQs on a website).
- Perishability means that an operation's output is difficult to store and ceases to have value after a relatively short time, so matching capacity with demand (or vice versa) is important to avoid either underutilised resources or lost revenue. Accurate demand forecasting helps to plan operations capacity.

Services and products are merging (and changing)

Increasingly the distinction between services and products is seen as not particularly useful. Some authorities see the essential purpose of all businesses, and therefore all operations, as being to 'serve customers'. Therefore, they argue, all operations are service providers who may (or may not) produce physical products as a means of serving their customers. This idea, that all operations should be seen as offering 'value propositions' through service, is called 'service-dominant logic'.[3] Among other things, it holds that service is the fundamental basis of exchange, that physical goods are simply the distribution mechanisms for the provision of service, and that the customer is always the co-creator of value. Our approach in this text is close to this in that we treat operations and process management as being important for all organisations. Whether they see themselves as manufacturers or service providers is very much a secondary issue.

Customers are part of the process – co-creation

If all operations can be seen as producing services, and services act on customers or their surrogates, then the role of customers on an operation's output should be considered. This is not a new idea, nor is it unusual for customers to play a central part in how they derive value from an operation's outputs. They browse the supermarket shelves, pick the items, take them to the checkout, sometimes scan them, then pack them and transport them home. They could reduce the extent of their participation by using a home delivery service that lets the retailer pick, pack and deliver their goods, though customers still have to be involved in ordering online. Similarly, patients visiting the doctor with an ailment are required to describe their symptoms and discuss alternative treatments; the better they can do this, the better the value they derive.

This idea of customer involvement is important because, in effect, the distinction between the traditional roles of 'producer' and 'consumer' are blurred. The concept is usually known either as co-creation or co-production; there is some disagreement as to what the two terms mean. Here we use the term 'co-creation' to describe all types of customer involvement, engagement, participation or collaboration with the operation, in one or all of the stages of the creation process. This has important implications for all operations, whether they see themselves as creating products, services or both. Not only does it emphasise the importance of customers in shaping how an operation's outputs can create value, it establishes the importance of a full two-way and transparent interaction between an operation and its customers.

Servitisation

Further evidence of the blending of services and products is the concept of 'servitisation'. This is the generic term that has come to mean any strategy that seeks to change the way in which product functionality is delivered to its markets. It means changing the focus of what an operation produces from its products to those services that the product enables. For example, one of the earliest, and most discussed, examples of servitisation comes from Rolls-Royce, the aero-engine manufacturer. Instead of selling its engines to its customers (mainly airlines), it offered them the option of paying a price per engine flying or operating hour, the price being dependent on the amount of support needed (for example, the number of aircraft, operating hours, amount of overhaul and repair, inventory, locations, transportation, etc.) and the performance guarantees required (for example, engine availability, turnaround time of repair, reliability, etc.). They originally called this service, 'power by the hour', a phrase seen by the industry as capturing the essence of servitisation.

Although originally dismissed by some as a management fad, it is an important indicator of how some operations are likely to develop. This is partly because it can generate extra revenue and partly because companies with a high installed base of capital goods (such as train manufacturers, construction vehicle makers, etc.) see their previous customers as representing a market that is worth exploiting. Partly also, services tend to provide a more stable source of revenue because they are resistant to the economic cycles that drive capital equipment purchases. In addition, servitisation can improve relationships with customers; after all, it is difficult to provide services for customers without getting to know them. In particular, offering services provides an opportunity to understand customer needs and motivates the translating of these needs through both product and service design. This is reflected in the case example, 'Servitisation and circular design at Philips lighting'.

Case example

Servitisation and circular design at Philips lighting[4]

Operations managers are increasingly having to re-evaluate how they think about their products and services and how they produce them. Take, for example, Philips Lighting,[5] who responded to developments in their markets by combining and adopting two important changes to operations practice – servitisation and the circular economy.

The company's servitisation offering is called 'lighting-as-a-service' (LaaS), where they take care of their customers' lighting needs, from the initial design and installation of the lighting to the operation and maintenance. By doing this, customers can save money because they pay only for the light they use, while at the same time avoiding the disturbance of having to replace and dispose of burnt-out bulbs or having to navigate system upgrades. The company originally became interested in LaaS when the architect Thomas Rau worked with Philips

Lighting to supply a novel 'pay-per-lux' intelligent lighting system, which was customised to fit the requirements of the Amsterdam office space of RAU Architects, while also reducing the price. When considering their lighting needs, Rau wanted to avoid buying an expensive over-engineered lighting system, only to eventually have to dispose of and replace it. Instead, he would rather purchase just the right amount of light 'as a service' that would suit the building. RAU and Philips developed a system that created a minimalist light plan, making as much use as possible of the building's natural sunlight. It combined a sensor and controller system that helped keep energy use to an absolute minimum, by darkening or brightening the artificial lighting in response to motion within a space or the presence of daylight. From the customer's point of view, they not only save money by paying only for the light they use, they also find it easier to optimise their use of energy, while avoiding the effort of managing the system. From the supplier's point of view, the agreement allowed Philips to retain control over how the lighting system worked, what products were supplied, how the system was maintained, how it was reconditioned and eventually how its products were recycled.

Other examples of Philips 'lighting-as-a-service' based on circular economy principles included the lighting for the terminal buildings at Amsterdam Airport Schiphol. This agreement means that Schiphol pays only for the light it uses, while Philips remains the owner of all fixtures and installations, responsible for the performance and durability of the system, and eventually its reuse and recycling at end of its useful life. Energy-efficient LED lamps resulted in a 50 per cent reduction in electricity consumption compared to conventional lighting systems, and the specially designed lighting fixtures were designed to last 75 per cent longer. Moreover, the collaboration between the supplier and user of the service resulted in reduced maintenance costs because components could be individually replaced instead of the entire fixture being recycled, resulting in significant raw material consumption and contributing to Schiphol's ambitious sustainability targets.

The process perspective at the level of the supply network

Any individual operation is part of a greater network of operations. It will have operations that supply it with the products and services it needs to create its own products and services. And unless it deals directly with the end consumer, it will supply its customer operations, who themselves will go on to supply their own customers. Moreover, any operation will probably have several suppliers, several customers and be in competition with other operations with whom it shares some suppliers and some customers. This collection of interconnected operations is called the supply network.

There are three important issues to understand about supply networks. First, they can be complex. Operations may have a large number of customers and suppliers who themselves have large numbers of customers and suppliers. The relationships between operations in the supply network can be subtle. One operation may be in direct competition with another in some markets while at the same time acting as collaborators or suppliers to each other in others. Second, theoretically the boundaries of any operation's supply network can be very wide indeed. They could go back to the operation that digs raw material out of the ground and go forward to the ultimate reuse and/or disposal of a product. Sometimes it is necessary to consider these ultimate boundaries (for example, when studying the environmental sustainability of products), but generally some kind of boundary to the network needs to be set so that more attention can be given to the more immediate operations in the network. Third, supply networks are always changing. Not only do operations sometimes lose customers and win others, or change their suppliers, they also may acquire operations that once were their customers or suppliers, or sell parts of their business, thus converting them into customers or suppliers.

Thinking about operations management in a supply network context is important. The overarching question for any operations manager is, 'Does my operation make a contribution to the supply network as a whole?' In other words, 'are we a good customer to our suppliers in the sense that the long-term cost of supply to us is reduced because we are easy to do business with?' And, 'are we good suppliers to our customers in the sense that, because of our understanding of the supply network as a whole, we understand their needs and have developed the capability to satisfy them?' Because of the significance of the supply network perspective, we deal with it twice more in this text: at a strategic level in Chapter 4, where we discuss the overall structure and scope of individual operations; and at a more operational level in Chapter 7, where we examine the role of supply chains in the delivery of products and services.

The process perspective at the level of the individual process

Because processes are smaller versions of operations, they have customers and suppliers in the same way as whole operations. We can view any operation as a network of individual processes that interact with each other, with each process being, at the same time, an internal supplier and an internal customer for other processes. This 'internal customer' concept provides a model to analyse the internal activities of an operation. If the whole operation is not working as it should, we may be able to trace the problem back along this internal network of customers and suppliers. It can also be a useful reminder to all parts of the operation that, by treating their internal customers with the same degree of care that they exercise on their external customers, the effectiveness of the whole operation can be improved. Again, remember that many of an organisation's processes are not operations processes, but are part of some other function. Table 1.2 illustrates just some of the processes that are contained within some of the more common non-operations functions, the outputs from these processes and their 'customers'.

There is an important implication of visualising each function of an organisation as being a network of processes. The diverse parts of a business are connected by the relationships between its various processes, and the organisational boundaries between each function and part of the business is really a secondary issue. Firms are always reorganising the boundaries between processes; they frequently move responsibility for tasks between departments. The tasks and the processes that perform them change less often. Similarly, tasks and processes may move between various businesses; that is what outsourcing and the 'do or buy' decision is all about (see Chapter 3). In other words, not only separate businesses but also whole supply networks can be seen as networks of processes. Who owns which processes and how the organisational boundaries between them move are separate decisions.

> **OPERATIONS PRINCIPLE**
> ***Whole businesses, and even whole supply networks, can be viewed as networks of processes.***

Table 1.2 Some examples of processes in non-operations functions

Organisational function	*Some of its processes*	*Some outputs from its process*	*Customer(s) for its outputs*
Marketing and sales	Planning process Forecasting process Order-taking process	Marketing plans Sales forecasts Confirmed orders	Senior management Sales staff, planners, operations Operations, finance
Finance and accounting	Budgeting process Capital approval processes Invoicing processes	Budget Capital request evaluations Invoices	Everyone Senior management, requestees External customers
Human resources management	Payroll processes Recruitment processes Training processes	Salary statements New hires Trained employees	Employees All other processes All other processes
Information technology	Systems review process Help desk process System implementation project processes	System evaluation Advice Implemented working systems and aftercare	All other processes All other processes All other processes

The 'line of sight' within process networks

Thinking about operations as networks of processes also gives a further benefit. It prompts the question of whether the people who operate each process have what is known as a clear 'line of sight' forward through to the external customers who will eventually be affected by their performance. If so, they will have a better chance of seeing how they contribute to the final value added for the operation's customers. Just as important, they will be better placed to help those other processes that lie between them and the customer to add value. Similarly, a clear 'line of sight' backwards helps to understand what is required from internal (and eventually external) suppliers. Certainly, a failure to understand how internal process chains work will reduce the effectiveness of the whole operation.

'End-to-end' business processes

It is worth remembering that what we choose to define as a specific process is not predetermined. We can define what is inside a process in any way we want. The boundaries between processes, the activities that they perform and the resources that they use, are all there because they have been designed in that way. It is common in an organisation to find processes defined by the type of activity engaged in. For example, invoicing processes, product design processes, sales processes, warehousing processes, assembly processes, painting processes, etc. This can be convenient because it groups similar resources together. But it is only one way of drawing the boundaries between processes. Theoretically, in large organisations there must be almost an infinite number of ways that activities and resources could be collected together as distinct processes. One way of redefining the boundaries and responsibilities of processes is to consider the 'end-to-end' set of activities that satisfy defined customer needs. Think about the various ways in which a business satisfies its customers. Many different activities and resources will probably contribute to 'producing' each of its products and services. Some authorities recommend grouping the activities and resources together in an end-to-end manner to satisfy each defined customer need. This approach is closely identified with the 'business process engineering' (or re-engineering) movement (examined in Chapter 12). It calls for a radical rethink of process design that will probably involve taking activities and resources out of different functions and placing them together to meet customer needs. Remember though, designing processes around end-to-end customer needs is only one way (although often the sensible one) of designing processes.

OPERATIONS PRINCIPLE
Processes are defined by how the organisation chooses to draw process boundaries.

Worked example

The 'Studio Division' (SD)

A broadcasting company has a 'Studio Division' (SD) that makes programmes for itself and other media companies. Original ideas usually come from the clients, although SD does share in the creative input. The business is described at the three levels of analysis in Figure 1.9.

At the level of the operation – SD produces media files, but its real output is the creativity and 'artistry' that is captured in the programmes. *'We interpret the client's needs and transform them into appealing and appropriate shows. We can do this because of the skills, experience and creativity of our staff, and our state-of-the-art technology,'* says the division's boss.

At the level of the supply network – the division has positioned itself to specialise in certain types of work. *'We did this so that we could develop a high level of expertise in a few relatively high-growth areas. It also reduces our*

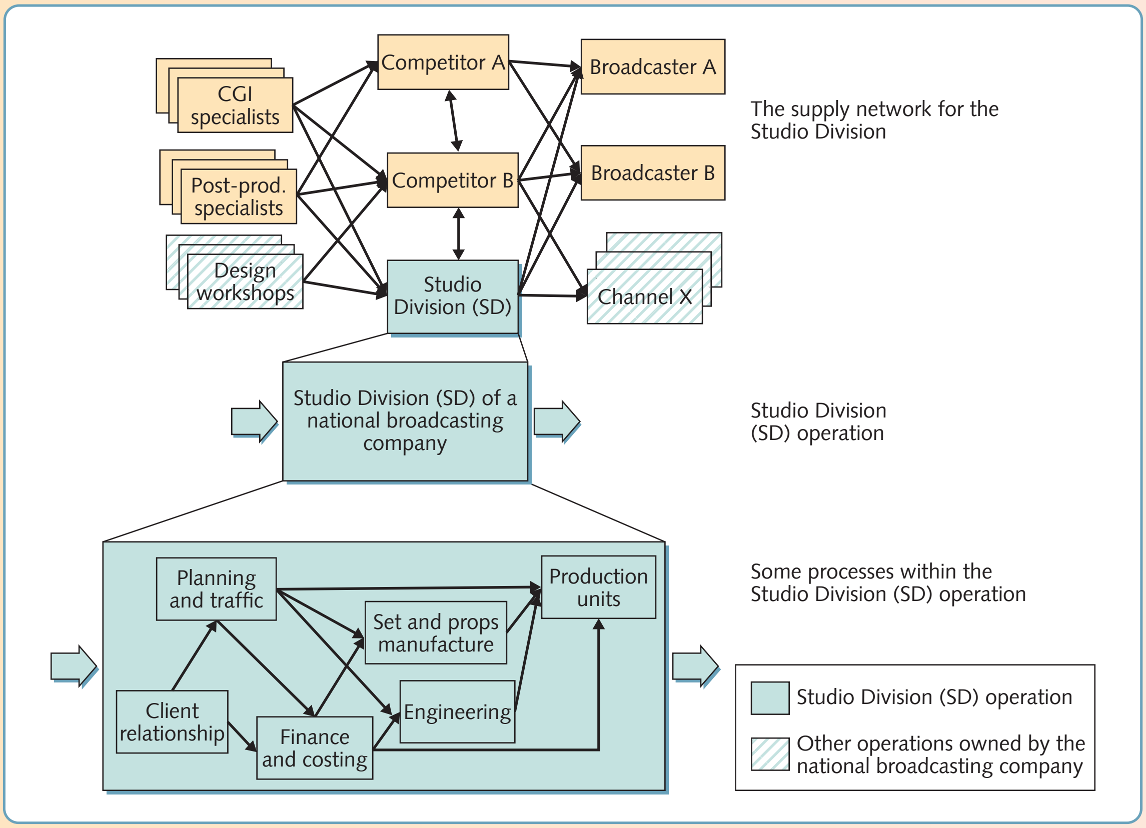


Figure 1.9 Operations and process management analysis for the Studio Division (SD) at three levels – the supply network, the operation and individual processes

dependence on our own broadcasting channels. Specialisation has also allowed us to outsource some activities such as computer-generated imagery imaging (CGI) and post-production that are no longer worth keeping in-house. However, our design workshop became so successful that they were "spun out" as a division in their own right and now work for other companies as well as ourselves.'

At the level of individual processes – many smaller processes contribute directly or indirectly to the production of programmes, including 'planning and traffic' who act as the operations managers for the whole operation, workshops that manufacture some of the sets, scenery and props for the productions, 'client liaison' who liaise with potential customers, 'engineering' that cares for and designs technical equipment, 'production units' that organise and shoot the programmes and 'finance and costing' that sets and controls operational budgets.

Creating end-to-end processes – SD produces products and services that fulfil customer needs. Each of these involves several of the existing departments within the company. For example, preparing a 'pitch' mainly needs the contributions of client relations and the finance and costing departments, but also needs smaller contributions from other departments. Figure 1.10 illustrates the contribution of each department to each product or service. The contributions of each department may not all occur in the same order. Currently, all the division's processes are clustered into conventional departments defined by the type of activity they perform: engineering, client relations, etc. A radical redesign of the operation could involve regrouping activities and resources into five 'business' processes that fulfil each of the five defined customer needs. This is shown diagrammatically by the dotted lines in Figure 1.10. It would involve the physical movement of resources (people and facilities) out of the current functional processes into the new end-to-end business processes. It is an example of how processes can be designed in ways that do not necessarily reflect conventional functional groupings.

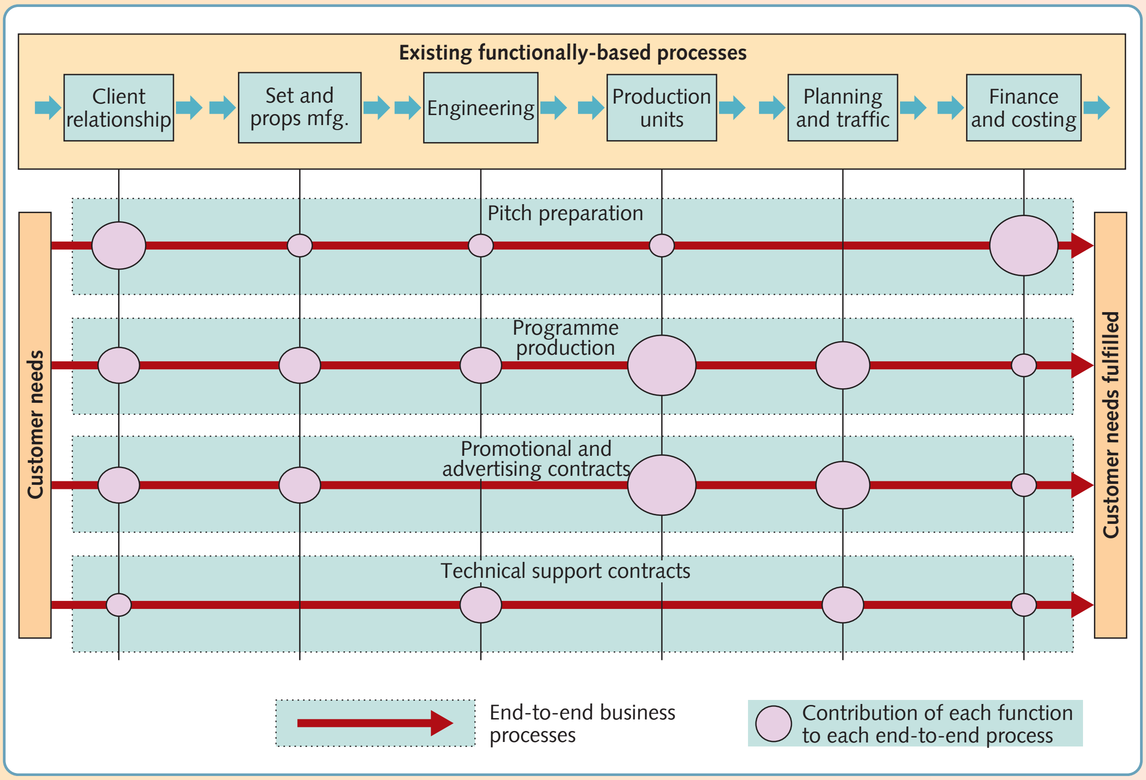


Figure 1.10 An example of how processes in the Studio Division (SD) could be reorganised around end-to-end business processes that fulfil defined customer needs

1.3 Diagnostic question: *Does operations and process management have a strategic impact?*

One of the biggest mistakes a business can make is to confuse 'operations' with 'operation*al*'. Operational is the opposite of strategic; it means detailed, localised, short-term, day-to-day. 'Operations', on the other hand, is the set of resources that produce products and services. Operations can be treated both at an operational *and a strategic level*. We shall examine some views of operations strategy in the next chapter. For now, we treat a fundamental question for any operation – how does the way we manage operations and processes have a strategic impact?

The operations function of an organisation is clearly significant strategically, if only because it often represents the bulk of its assets and the majority of its people. Yet its true value is more than 'bulk'. It can 'make' the business in the sense that it gives the ability to compete, through both the short-term ability to respond to customers and the long-term capabilities that will keep it ahead of its competitors. But if an operations function cannot produce its products and services effectively, it could 'break' the business by handicapping its performance no matter how it positions and sells itself in its markets. In addition, all the processes within an organisation, irrespective of which function they support, can help (or hinder) it in achieving its strategic objectives.

Three levels of operations performance

Operations performance can be assessed at three different levels. At its broadest, operations can be judged on how it impacts on long-term societal issues. At the level of the individual firm or enterprise, it can be judged on how it supports (or not) their strategic aims. At its more operational level, the individual processes within an operation can be judged by how well they serve their (internal and external) customers and improve the efficiency with which they do it. In the rest of this section we will look at each of these three levels.

Operations performance at a societal level

Operations decisions affect a whole variety of 'stakeholders' – the people and groups that have a legitimate interest in the operation's activities. These include the operation's employees, its customers, its shareholders, its immediate community and society in general. But although each of these groups will be interested in an operation's performance, they are likely to have very different views of which aspect of performance is important. Nevertheless, if one is to judge operations at a broad societal level, one must judge the impact it has on its stakeholders.

The triple bottom line

One idea that tries to capture the idea of a broader approach to assessing an organisation's performance is the 'triple bottom line'[6] (TBL, or 3BL), also known as 'people, planet and profit'. Essentially, it is a straightforward idea simply that organisations should measure themselves not just on the traditional economic profit that they generate for their owners, but also on the impact their operations have on society (broadly, in the sense of communities, and individually, for example in terms of their employees) and the ecological impact on the environment. The influential initiative that has come out of this triple bottom line approach is that of 'sustainability'. A sustainable business is one that creates an acceptable profit for its owners, but minimises the damage to the environment and enhances the existence of the people with whom it has contact. In other words, it balances economic, environmental and societal interests. This gives the organisation its 'license to operate' in society. The assumption underlying the triple bottom line (which is not universally accepted, see the critical commentary at the end of the chapter) is that a sustainable business is more likely to remain successful in the long term than one that focuses on economic goals alone.

> **OPERATIONS PRINCIPLE**
> ***Operations should judge themselves on the triple bottom line principle of 'people, planet and profit'.***

However, the idea of triple bottom line performance does have its critics. It is unreasonable, some say, for operations managers to have to cope with the conflicting pressures of maximising profitability as well as managing in the interests of society in general with accountability and transparency. Even if a business wanted to reflect aspects of performance beyond its own immediate interests, how is it to do that? According to Michael Jensen of Harvard Business School, '*At the economy-wide or social level, the issue is this: If we could dictate the criterion or objective function to be maximized by firms (and thus the performance criterion by which corporate executives choose among alternative policy options), what would it be? Or, to put the issue even more simply: How do we want the firms in our economy to measure their own performance? How do we want them to determine what is better versus worse?*'[7] He also holds that using stakeholder perspectives gives undue weight to narrow special interests who want to use the organisation's resources for their own ends. The stakeholder perspective gives them a spurious legitimacy that '*undermines the foundations of value-seeking behaviour*'.

Operations performance at the level of the enterprise

The ability of operations and process management to impact the strategic success of any kind of enterprise is being increasingly recognised. When compared with only a few years ago, it now attracts far more attention and, according to some reports, accounts for the largest share of all the money spent by businesses on consultancy advice. This may be partly because the area has been neglected in the past. But it also denotes an acceptance that it can have both short-term and long-term impact. This can be seen in the impact that operations and process management can have on the business's costs, revenue, risk, investment and capabilities:

- It can reduce the *costs* of producing products and services by being efficient. The more productive the operation is at transforming inputs into outputs, the lower will be the cost of producing a unit of output. Cost is never totally unimportant for any business, but generally the higher the cost of a product or service when compared to the price it commands in the market, the more important cost reduction will be as an operations objective. Even so, cost reduction is almost always treated as an important contribution that operations can make to the success of any business.
- It can increase *revenue* by increasing customer satisfaction through quality, service and innovation. Existing customers are more likely to be retained and new customers are more likely to be attracted to products and services if they are error-free and appropriately designed, if the operation is fast and responsive in meeting their needs and keeping its delivery promises, and if an operation can be flexible, both in customising its products and services and introducing new ones. It is operations that directly influence the quality, speed, dependability and flexibility of the business, all of which have a major impact on a company's ability to maximise its revenue.
- It can reduce the *risk* of operational failure, because well-designed and well-run operations should be less likely to fail. All failures can eventually be traced back to some kind of failure within a process. Furthermore, a well-designed process, if it does fail, should be able to recover faster and with less disruption (this is called *resilience*).
- It can ensure *effective investment* (*capital employed)* to produce its products and services. Eventually all businesses in the commercial world are judged by the return that they produce for their investors. This is a function of profit (the difference between costs and revenues) and the amount of money invested in the business's operations resources. We have already established that effective and efficient operations can reduce costs and increase revenue. What is sometimes overlooked is the role of operations in reducing the investment required per unit of output. It does this by increasing the effective capacity of the operation and by being innovative in how it uses its physical resources.
- It can *build capabilities* that will form the basis for *future* innovation by building a solid base of operations skills and knowledge within the business. Every time an operation produces a product or a service it has the opportunity to accumulate knowledge about how that product or service is best produced. This accumulation of knowledge should be used as a basis for learning and improvement. If so, in the long term, capabilities can be built that will allow the operation to respond to future market challenges. Conversely, if an operations function is simply seen as the mechanical and routine fulfilment of customer requests, then it is difficult to build the knowledge base that will allow future innovation.

> **OPERATIONS PRINCIPLE**
> ***All operations should be expected to contribute to their enterprise by controlling costs, increasing revenue, reducing risk, making investment more effective and growing long-term capabilities.***

Operations performance at the level of an operation's processes

Operations and process management can also be judged at the more operational level. This is how well the network of processes within an operation serves its internal, and eventually its external, customers. There are five aspects of operations and process performance, all of which to a greater or lesser extent will affect customer satisfaction and business competitiveness:

1. ***Quality*** – doing things right, providing error-free goods and services that are 'fit for their purpose'.
2. ***Speed*** – doing things fast, minimising the time between a customer asking for goods and services and the customer receiving them in full.
3. ***Dependability*** – doing things on-time, keeping the delivery promises that have been made to customers.
4. ***Flexibility*** – changing what you do or how you do it, the ability to vary or adapt the operation's activities to cope with unexpected circumstances or to give customers individual treatment, or to introduce new products or services.
5. ***Cost*** – doing things cheaply, producing goods and services at a cost that enables them to be priced appropriately for the market while still allowing a return to the organisation (or, in a not-for-profit organisation, that give good value to the taxpayers or whoever is funding the operation).

> OPERATIONS PRINCIPLE
> ***Operations and process performance at an operational level can be grouped together as quality, speed, dependability, flexibility and cost.***

But do not think that these 'operational' aspects of performance have little strategic significance. On the contrary, they all contribute to the success of the organisation as a whole. We will look further at these aspects of performance in the next chapter, which deals with operations strategy.

Worked example

The 'Studio Division' (SD) (continued)

The SD, described earlier, should be able to identify all four ways in which its operations and processes can have a strategic impact. The division is expected to generate reasonable returns by controlling its costs and being able to command relatively high fees. *'Sure, we need to keep our costs down. We always review our budgets for bought-in materials and services. Just as important, we measure the efficiency of all our processes, and we expect annual improvements in process efficiency to compensate for any increases in input costs.'* (Reducing costs.) *'Our services are in demand by customers because we are good to work with,'* says the division's managing director. *'We have the technical resources to do a really great job and we always give good service. Projects are completed on time and within budget. More importantly, our clients know that we can work with them to ensure a high level of programme creativity. We are not the cheapest, but we are the best.'* (Increasing revenue.) *'Also, we have a robust set of processes that minimise the chances of projects failing to achieve success.'* (Reducing risk.) The division also has to justify its annual spend on equipment to its main board. *'We try and keep up to date with the new technology that can really make an impact on our programme making, but we always have to demonstrate how it will improve profitability.'* (Effective investment.) *'We also try to adapt new technology and integrate it into our creative processes in some way so that gives us some kind of advantage over our competitors.'* (Build capabilities.)

Case example

MSF operations provide medical aid to people in danger[8]

Do not assume that the operations and process perspective applies only to for-profit enterprises. Terms such as *competitive advantage*, *markets* and *business* that are used in this text are usually associated with companies in the for-profit sector. Yet operations management is also relevant to organisations whose purpose is not primarily to earn profits. For example, Médecins Sans Frontières, MSF (Doctors Without Borders), the independent humanitarian organisation, uses its operations processes to provide medical aid where it is most needed in countries around the world, usually in crisis situations such as armed conflicts, epidemics, famines and natural disasters. Its teams deliver both medical aid and material aid quickly and efficiently. It is one of the most admired and effective relief organisations in the world. But no amount of fine intentions can translate into effective action without superior operations management. As MSF says, it must be able to react to any crisis with 'fast response, efficient logistics systems, and efficient project management'. Its response process has five phases: proposal, assessment, initiation, running the project and closing. The information that prompts a possible mission can come from governments, the international community, humanitarian organisations or MSF teams already present in the region. Once the information has been checked and validated, MSF sends a team of medical and logistics experts to the crisis area to carry out a quick evaluation. The team assesses the situation, the number of people affected, and the current and future needs, and sends a proposal back to the MSF office. When the proposal is approved, MSF staff starts the process of selecting personnel, organising materials and resources and securing project funds. Initiating a project involves sending technical equipment and resources to the area. In large crises, planes fly-in all the necessary materials so that the work can begin immediately. Thanks to their pre-planned processes, specialised kits and the emergency stores, MSF can distribute materials and equipment within 48 hours, ready for the response team to start work as soon as they arrive. Most MSF projects generally run for somewhere between 18 months and three and a half years. Whether an emergency response or a long-term healthcare project, the closing process is roughly similar. Once the critical medical needs have been met (which could be after weeks, months or years, depending on the situation), MSF begins to close the project with a gradual withdrawal of staff and equipment. At this stage, the project closes or is passed on to an appropriate organisation. MSF will also close a project if risks in the area become too great to ensure staff safety.

1.4 Diagnostic question: *Are processes managed to reflect their operating circumstances?*

OPERATIONS PRINCIPLE
The way in which processes need to be managed is influenced by volume, variety, variation in demand and visibility.

All processes differ in some way, so, to some extent, all processes will need to be managed differently. Some of the differences between processes are 'technical' in the sense that different products and services require different skills and technologies to produce them. However, processes also differ in terms of the nature of demand for their products or services. Four characteristics in particular have a significant effect on how processes need to be managed:

- The *volume* of the products and services produced.
- The *variety* of the different products and services produced.
- The *variation in demand* for products and services.
- The *degree of visibility* that customers have of the production of products and services.

Volume – Processes with a high volume of output will have a high degree of repeatability, and because tasks are repeated frequently it often makes sense for staff to specialise in the tasks they perform. This allows the systemisation of activities, where standard procedures may be codified and set down in a manual with instructions on how each part of the job should be performed. Also, because tasks are systemised and repeated, it is often worthwhile developing specialised technology that gives higher processing efficiencies. By contrast, low-volume processes with less repetition cannot specialise to the same degree. Staff is likely to perform a wide range of tasks, and while this may be more rewarding, it is less open to systemisation. Nor is it likely that efficient, high-throughput technology could be used. The implications of this are that high-volume processes have more opportunities to produce products or services at low unit cost. So, for example, the volume and standardisation of large fast-food restaurant chains, such as McDonald's or KFC, enables them to produce with greater efficiency than a small, local cafeteria or diner.

Variety – Processes that produce a high variety of products and services must engage in a wide range of different activities, changing relatively frequently between each activity. They must also contain a wide range of skills and technology sufficiently 'general purpose' to cope with the range of activities and sufficiently flexible to change between them. A high level of variety may also imply a relatively wide range of inputs to the process and the additional complexity of matching customer requirements to appropriate products or services. High-variety processes are invariably more complex and costly than low-variety ones. For example, a taxi company is usually prepared to pick up and drive customers almost anywhere (at a price); they may even take you by the route of your choice. There are an infinite number of potential routes (products) that it offers. But, its cost per kilometre travelled will be higher than a less-customised form of transport such as a bus service.

Variation in demand – Processes are generally easier to manage when they only have to cope with predictably constant demand: resources can be geared to a level that is just capable of meeting demand, and all activities can be planned in advance. By contrast, when demand is variable and/or unpredictable, resources will have to be adjusted over time. Worse still, when demand is unpredictable, extra resources will have to be designed into the process to provide a 'capacity cushion' that can absorb unexpected demand. So, for example, processes that manufacture high-fashion garments will have to cope with the general seasonality of the garment market, together with the uncertainty of whether particular styles may or may not prove popular. Operations that make conventional business suits are likely to have less fluctuation in demand over time, and be less prone to unexpected fluctuations. Because processes with lower variation do not need any extra safety capacity and can be planned in advance, they will generally have lower costs than those with higher variation.

Visibility – Process visibility is a slightly more difficult concept to envisage. It indicates how much of the process is 'experienced' directly by customers, or how much the process is 'exposed' to its customers. Generally, processes that act directly on customers (such as retail processes or healthcare processes) will have more of their activities visible to their customers than those that act on materials and information. However, even material- and information-transforming processes may provide a degree of visibility to the customers. For example, parcel distribution operations provide internet-based 'track and trace' facilities to enable their customers to have visibility of where their packages are at any time. Low-visibility processes, if they communicate with their customers at all, do so using less immediate channels such as the telephone or the internet. Much of the process can be more 'factory-like'. The time lag between customer request and response could be measured in days rather than the near-immediate response expected from high-visibility processes. This lag allows the activities in a low-visibility process to be performed when it is convenient to the operation, so achieving high utilisation. Also, because the customer interface needs managing, staff in high-visibility processes need customer contact skills that shape the customer's perception of process performance. For all these reasons, high-visibility processes tend to have higher costs than low-visibility processes.

Many operations have both high- and low-visibility processes. This serves to emphasise the difference that the degree of visibility makes. For example, in an airport, some of its processes are relatively visible to its customers (check-in desks, the information desks, restaurants, passport control and security staff, etc.). These staff operate in a high-visibility 'front-office' environment. Other processes in the airport have relatively little, if any, customer visibility (baggage-handling processes, overnight freight operations, loading meals on to the aircraft, cleaning, etc.). We rarely see these processes; they perform the vital but low-visibility tasks, in the 'back-office' part of the operation.

The implications of the four Vs of processes

All four dimensions have implications for processing costs. Put simply, high volume, low variety, low variation and low visibility all help to keep processing costs down. Conversely, low volume, high variety, high variation and high customer contact generally carry some kind of cost penalty for the process. This is why the volume dimension is drawn with its 'low' end at the left, unlike the other dimensions, to keep all the 'low cost' implications on the right. Figure 1.11 summarises the implications of such positioning. The four dimensions also have implications

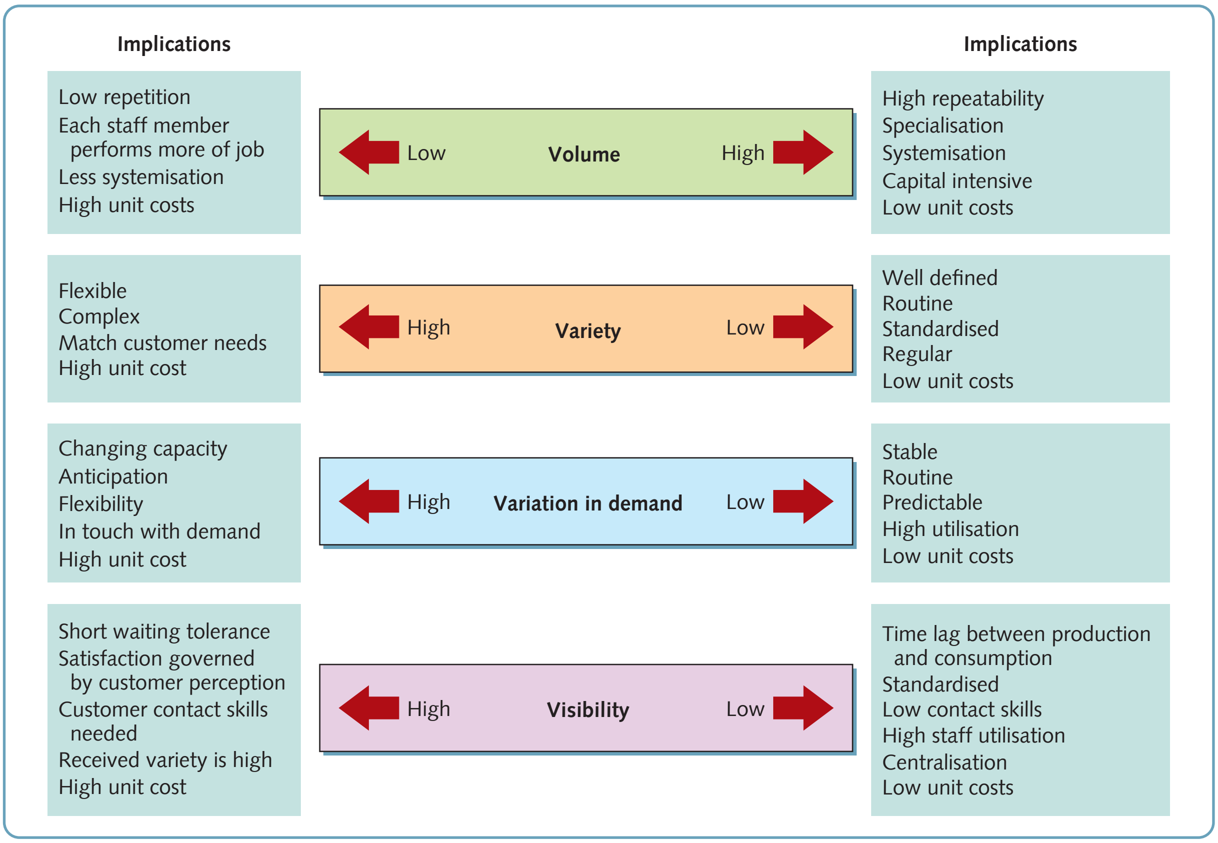


Figure 1.11 The implications of the four Vs of processes

for the types of activities that operations managers will have to focus on. As we will see in later chapters, volume and variety are particularly important in how processes are resourced and designed (Chapters 4 and 5), variation in demand calls for careful consideration of how capacity is to be managed (Chapter 8), while visibility has a significant influence over how the quality of products and services is judged (Chapter 13).

Charting processes using the four Vs

In almost any operation, processes can be identified that have different positions on the four dimensions, and which therefore have different objectives and will need managing in different ways. To a large extent the position of a process on the four dimensions is determined by the demand of the market it is serving. However, most processes have some discretion in moving themselves on the dimensions. Look at the different positions on the visibility dimension that retail banks have adopted. At one time, using branch tellers was the only way customers could contact a bank. Now access to the bank's services could be through (in decreasing order of visibility) a personal banker, who visits your home or office, a conversation with a branch manager, the teller at the window, telephone contact through a call centre, internet banking services or an ATM cash machine. These other processes offer services that have been developed by banks to serve different market needs.

OPERATIONS PRINCIPLE
Operations and processes can (other things being equal) reduce their costs by increasing volume, reducing variety, reducing variation and reducing visibility.

Figure 1.12 illustrates the different positions on the four Vs for some retail banking processes. Note that the personal banking/advice service is positioned at the high-cost end of the four Vs. For this reason, such services are often offered only to relatively wealthy customers that represent high profit opportunities for the bank. Note also that the more recent developments in retail banking, such as call centres, internet banking and ATMs, all represent a shift towards the low-cost end of the four Vs. New processes that exploit new technologies can often have a profound impact on the implications of each dimension. For example, internet banking, when compared with an ATM cash machine, offers a far higher variety of options for customers, but because the process is automated through its information technology, the cost of offering this variety is less than at a conventional branch or even a call centre.

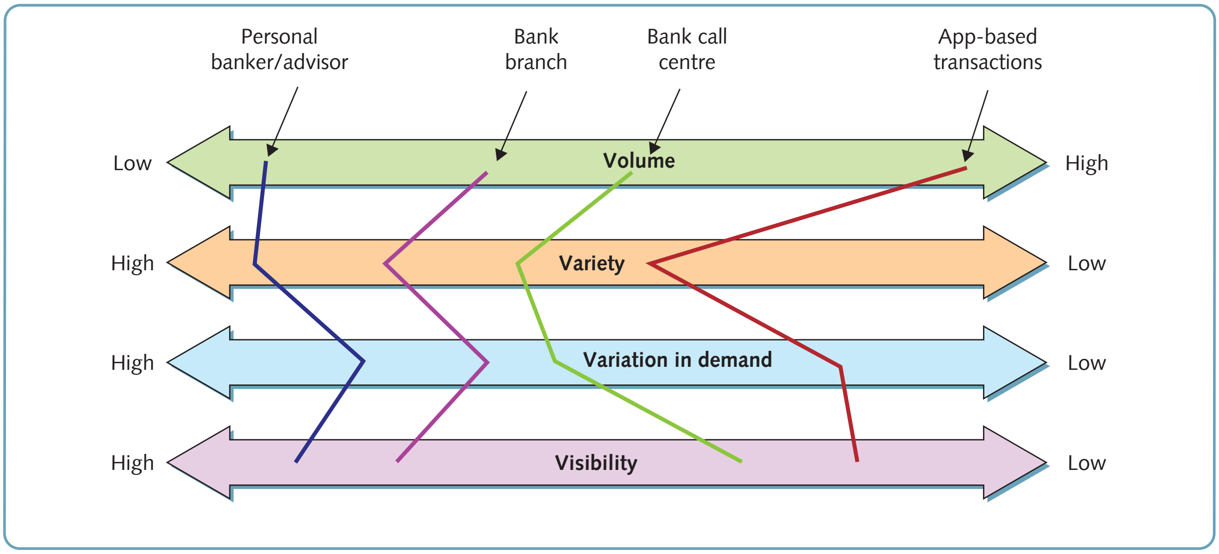


Figure 1.12 **Four Vs analysis for some retail banking processes**

1.5 Diagnostic question: *Is operations and process decision-making appropriate?*

> **OPERATIONS PRINCIPLE**
> ***Operations management activities can be grouped into four broad categories: directing the overall strategy of the operation; designing the operation's products, services and processes; planning and controlling delivery; and developing process performance.***

Managing operations and processes involves a whole range of separate decisions that will determine how well they achieve their overall purpose and contribute to the organisation as a whole. The way operations managers approach decision-making is therefore of considerable importance in determining their effectiveness. But different operations decisions serve different purposes. They can be grouped together in various ways. Look at other texts on operations management and you will find many different ways of structuring operations decisions and therefore the subject as a whole. Here we have chosen to classify activities into four broad groups, relating to four broad activities that, more or less, follow a sequence that corresponds to the life cycle of operations and processes:

- *Directing* the overall strategy of the operation. A general understanding of operations and processes and their strategic purpose, together with an appreciation of how strategic purpose is translated into reality through how innovation is incorporated into products and services, and how much of the total value-adding process should be kept in-house and how much outsourced.
- *Designing* the operation's processes. Design is the activity of determining the physical form, shape and composition of operations and processes, together with the types of resources they contain.
- Planning and controlling process *delivery*. After being designed, the delivery of products and services from suppliers and through the total operation to customers must be planned and controlled.
- *Developing* process performance. Increasingly it is recognised that operations managers, or any process managers, cannot simply deliver products and services routinely in the same way that they always have done. They have a responsibility to develop the capabilities of their processes to improve process performance.

A model of operations and process management

We can now combine two ideas to develop the model of operations and process management that will be used throughout this text. The first is the idea that *operations* and the *processes* that make up both the operations and other business functions are transformation systems that take in inputs and use process resources to transform them into outputs. The second idea is that the resources both in an organisation's operations as a whole and in its individual processes need to be managed in terms of how they are *directed*, how they are *designed*, how *delivery* is planned and controlled and how they are *developed* and improved. Figure 1.13 shows how these two ideas go together. This text will use this classification to examine the more important decisions that should be of interest to all managers of operations and processes.

However, although it is useful to categorise operations decisions in this way, the boundaries between these four categories are not 'clean'. There are always overlaps and interrelationships between the categories. The decisions that we classify as 'directing' will, of course, impact on all other operations decisions – that is the definition of 'directing'. And how processes are designed will limit how delivery can be organised and how easy it is to develop their capabilities. 'Delivery' decisions, such as whether to adopt lean practices, very much impact on how improvement (development) happens, and so on. All operations decisions are interrelated. Our

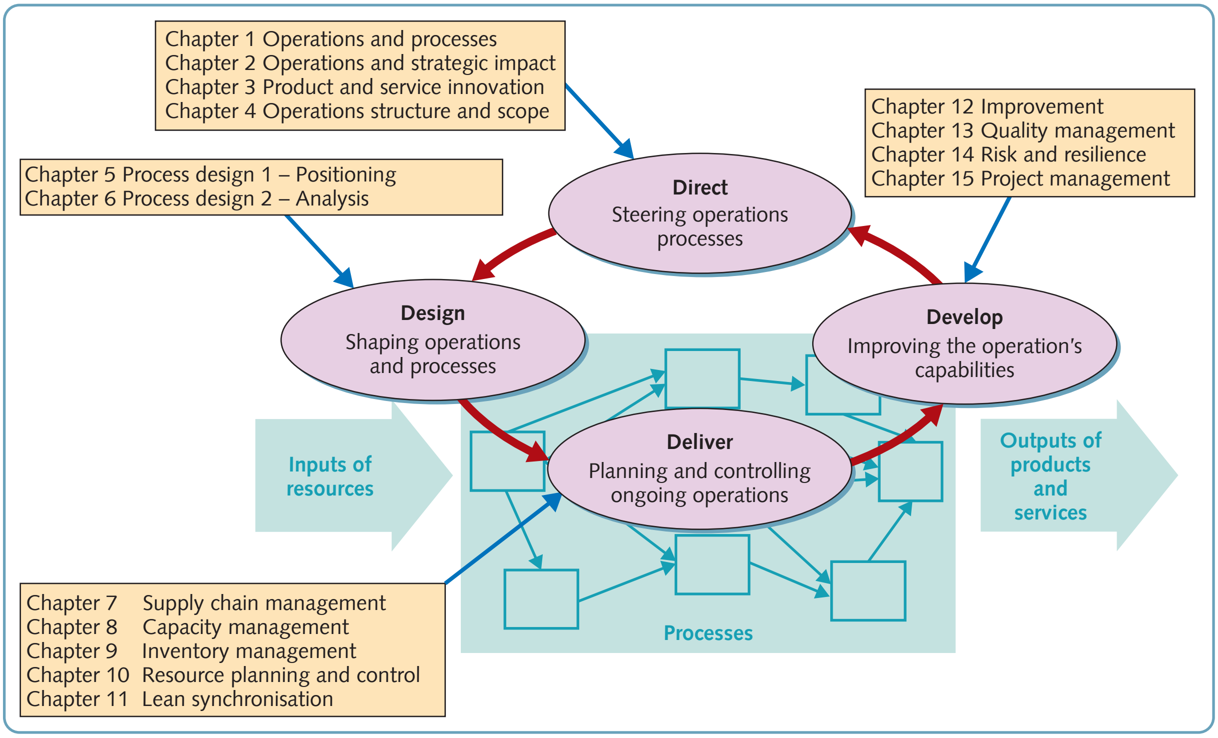


Figure 1.13 **A classification model of operations and processes management decisions**

categorisation of decisions simply indicates their main purpose. Moreover, such decisions are not always obvious and depend on 'technical' assumptions. Note how the case example 'Adidas shuts its robotic factories' demonstrates how one company, although being innovative in tackling changes in market requirements, still had to modify its operations decisions.

Case example

Adidas shuts its robotic factories[9]

It sounded such a good idea. Although, like almost all of its rivals, Adidas had tended to concentrate on the design, marketing and distribution of its trainers (sneakers), subcontracting the 'making' part of the total process to a complex supply network located largely in Asia. It had not run or owned its own manufacturing operations since the 1990s. The network of suppliers it employed spread over more than 1,000 facilities in 63 countries. Yet, like other similar companies, Adidas faced some problems with its Asian outsourcing model. Growing affluence in the area had increased costs, the longer and complex supply networks were difficult to control, and such a globalised and complex supply chain meant a long lead-time (around 18 months) between conceiving a new trainer and it arriving in the shops. And it is this last point that was the most problematic, particularly for fashionable trainers with a short 'fashion life'. Even orders to replenish stocks can take two to three months. But fashion cycles for trainers are getting shorter, with some designs lasting only one to three years. Faced with this, Adidas developed its 'Speedfactory' operation, the first of which was located in Germany and the second in the US. The Speedfactory was totally automated, and designed to be able to accommodate new technologies, such as 3D printing enabled by motion-capture technology. And because almost all the stages of manufacturing were done on the same site, the intention was to make Adidas faster and more flexible, especially in producing small batches of fashionable products. It was hoped that the Speedfactories could produce shoes in days and replenish the fastest-selling products during the same season.

Yet within four years of the Speedfactories opening, Adidas announced it would cease production at the facilities. The company said it made more sense for the company to concentrate its Speedfactory production in Asia where the know-how and the vast majority of its suppliers were located, and where Adidas already makes more than 90 per cent of its products. Adidas said it would use its Speedfactory technology at two Asian supplier factories, and would concentrate on modernising its other suppliers. One reason for the relative failure of the Speedfactories was the restricted range of models they could make. Adidas had set up Speedfactories to make trainers with a knit upper and Adidas's unique bouncy 'Boost' midsole, but it could not make leather shoes with a rubber sole because that used a different kind of joining process. So, as was pointed out by commentators, the effort was a failure not because its objective was flawed, but rather because it paid insufficient attention to the manufacturing process itself. Moreover, as Adidas pointed out, the learning that it gained from the Speedfactories would be used in its Asian supply base.

Quantitative and qualitative decision-making in operations

In operations and process management (and throughout this text) we frequently use 'models', the better to understand a decision. By a model we mean an explicit statement of our image of reality. It is a representation of the aspects of a decision with which we are concerned. It structures and formalises the information we possess about the decision, and in doing so presents reality in a simplified and organised form. A model therefore provides an abstraction of a more complex reality. They can be partial in that they exclude some factors, and they can aggregate or compress several factors into one, but models are at the core of operations decision-making. Some of the models used in operations are qualitative. They categorise or describe the relationships between aspects of decisions, but they do not necessarily ascribe precise associations between variables. For example, Figure 1.9, which describes the relationships between processes and operations for the Studio Division, is a qualitative model. It does not provide an 'answer', but it does enhance understanding and stimulate discussion around other possible ways to organise the operations.

Quantitative models are also important in operations and process management, but present different challenges. Quantitative models try to represent the underlying behaviours involved in a decision by using mathematical and/or statistical descriptions of relationships. They allocate numerical values to variables to produce a mathematical representation of reality. For example, the economic order quantity (EOQ) model that we explain in Chapter 9 is a good illustration of a quantitative model. It gives a precise relationship between the costs involved in making one particular inventory decision and therefore can be used to make the decision of how much stock to order. Well, at least it is supposed to. In fact, this model illustrated one of the problems with using a quantitative approach in operations and process management: in order to model the decision mathematically, reality has to be simplified to an extent that may severely limit its usefulness. Not that this is a condemnation of quantitative modelling. Practical operations management depends on the quantification of decision-making where possible. But for most operations decisions some combination of quantitative and qualitative modelling is required.

'Behavioural' operations

Academics who write about, research or teach operations management are sometimes accused of ignoring the 'practical reality' of how operations management decisions are made. Their models, frameworks and guidance, it is claimed, do not reflect how people really behave when making operations management decisions in practice. This has led to the development of a (relatively) new branch of operations management. It is called 'behavioural operations management' (BOM) or simply 'behavioural operations', and explores the interaction of human behaviours and operational systems and processes. More specifically, it challenges the idea that managers are rational when making decisions that impact operations performance. One research team has summarised what they see as common behavioural assumptions to operations models:[10]

- People are not a major factor in operations decisions.
- People are deterministic and predictable.
- People make decisions independently of each other.
- People do not learn from experience.
- People are not part of the product or service.
- People are emotionless.
- Work is perfectly observable and can be understood.

Clearly these assumptions are extremely unrealistic, and while no experienced operations manager would ever subscribe to them, they do act as a warning as to how the models, frameworks and techniques in this text should *not* be interpreted. Very few of the models that we use are rigidly prescriptive. Generally, they do not attempt to dictate a single 'optimum' solution to any problem. Rather they try to structure and clarify operations decisions, the better to understand, debate and, hopefully, make better decisions.

Critical commentary

All chapters will contain a short critical commentary on the main ideas covered in the chapter. Its purpose is not to undermine the issues discussed in the chapter, but to emphasise that, although we present a relatively orthodox view of operations, there are other perspectives.

- The central idea in this introductory chapter is that all organisations have operations (and other functions) that have processes that produce products and services, and that all these processes are essentially similar. However, some believe that by even trying to characterise organisations in this way (perhaps by calling them 'processes') one loses or distorts their nature and depersonalises or takes the 'humanity' out of the way in which we think of the organisation. This point is often raised by 'professional' staff. For example, the head of one European 'Medical Association' (a doctors' trade union) criticised hospital authorities for expecting a 'sausage factory service based on productivity targets'. No matter how similar they appear on paper, it is argued, a hospital can never be viewed in the same way as a factory.
- At least some of this discomfort at the idea that the process perspective is demeaning in some way is the (false) assumption that 'people' and 'process' are different things. Yet a process is simply 'how you do things', or more formally, 'an arrangement of resources and activities that transform inputs into outputs', and one of the most important resources is the people who staff the process. People are integral to processes, and processes are what people do. However, there are dangers in taking an exclusively mechanistic process perspective. Notwithstanding the point we made earlier about the risk that a process perspective can depict the messy reality of organisations in a naïve manner, in our view it is a risk well worth taking.

SUMMARY CHECKLIST

Each chapter contains a summary checklist in the form of questions that can be usefully applied to any type of operation and reflect the major diagnostic questions used within the chapter.

- ☐ Is the operations function of the business clearly defined?
- ☐ Do operations managers realise that they are operations managers even if they are called by some other title?
- ☐ Do the non-operations functions within the business realise that they manage processes?
- ☐ Does everyone understand the inputs, activities and outputs of the processes of which they are part?
- ☐ Is the balance between products and services produced by the operations function well understood?
- ☐ Are changes that may occur in the balance between products and services produced by the operation understood?
- ☐ Does the operation assess its impact on the environment and society as well as its financial performance?
- ☐ What contribution is operations making towards reducing costs, increasing revenues, reducing the risks of failure, making better use of capital employed and developing the capability for future innovation?
- ☐ Does the operation understand its position in, and contribution to, the overall supply network?
- ☐ Are the individual processes that comprise the operations function defined and understood?
- ☐ Are individual processes aware of the internal customer and supplier concept?
- ☐ Do they use the ideas and principles of operations management to improve the performance of their processes?
- ☐ Has the concept of end-to-end business processes been examined and considered?
- ☐ Are the differences (in terms of volume, variety, variation and visibility) between processes understood?
- ☐ Are the volume, variety, variation and visibility characteristics of processes reflected in the way they are managed?
- ☐ Do all operations managers understand the full range of decisions that they should be involved in?
- ☐ Are all operations managers aware of the various qualitative and quantitative models that can help them in their decision-making?

Case study

Kaston-Trenton Service (KTS)

Kaston-Trenton Service (KTS) was a domestic heating boiler maintenance company, based in the East of England. Founded in the 1960s by plumber Christopher Trenton, it had grown substantially and was now run jointly by Christopher's two children, Ros, who looked after all marketing, sales and finance, and Mark, who looked after operations and supply issues. The company initially offered maintenance and repair services to domestic (household) customers with gas or oil-burning boilers and had expanded into offering similar services to business customers. Within the last two years KTS had also moved beyond simply servicing systems, to designing and installing HVAC (heating, ventilating and air conditioning) systems for business customers.

'*Expanding into the design and installation business was something of a gamble*', according to Ros. '*At the time, the B2B [business-to-business] part of our work was clearly showing more growth potential than our traditional domestic business, and servicing business customers was also more profitable. So far, the installation venture has had mixed success. The jobs that we have done have been successful and our new customers very satisfied, but so far we have lost money on them. Partly, this is because we have had to invest in extra workshop space at our headquarters and employ a system designer, who is relatively expensive (but good) and only partly utilised at the moment. Hopefully, profitability will improve as the volume of installation jobs increases.*'

Table 1.3 shows the number of contracts and the revenue from domestic servicing, business servicing and the design and installation businesses, both for the previous year and the forecast for the current year of operation (all figures as of end Qtr 3). The profitability of the three offerings was difficult to determine exactly, but Ros and Mark were satisfied with the contribution of domestic boiler servicing, and especially of the business boiler servicing activities.

KTS services

Domestic boiler servicing was seen by Ros and Mark as a 'cash cow', generating revenues at a fairly steady rate. There were many different makes of boiler installed, but KTS only contracted to service the most common, which accounted for about 60 per cent of the installed base. Less-common boilers were often serviced by the manufacturers that supplied them. Domestic servicing accounted for by far the most individual contracts for KTS, with customers spread over most of the East of England. Around 95 per cent of customers renewed their contracts each year, which was seen as a testament both to their quality of service and the company's keen pricing. '*It's a price-sensitive market*', said Ros. '*We have to be competitive, but that's not all that counts. Most visits by our technicians are routine yearly services, but about 20 per cent of visits are 'call-outs' with varying degrees of urgency. If a home boiler stops working on a winter weekend, the householder obviously expects us to respond quickly, and we try our best to get a technician to them within four or five hours. If it's simply a non-urgent controller fault in summer, we would probably agree a mutually convenient time to visit within a couple of days. Actually, the idea of a 'mutually convenient time' is important in this market. Householders often have to make special arrangements to be in, so we have to be flexible in arranging appointments and absolutely reliable in being there on time. Although call-outs are only 20 per cent of visits, they cause the majority of problems because both their timing and duration are unpredictable. Also, customers are sensitised to boiler performance following an emergency call-out. What we call the 'robustness of the repair' has to be high. Once it's fixed, it should stay fixed, at least for a reasonable length of time*'.

Business boiler servicing was different. Most customers' systems had been, to some extent, customised, so the

Table 1.3 The number of contracts and the revenue from the three activities

Activity	*Previous year*		*Current year (forecast)*	
	Number of contracts	*Revenue (£000)*	*Number of contracts*	*Revenue (£000)*
Domestic boiler servicing	7331	1408	9700	1930
Business boiler servicing	972	699	1354	1116
Design and installation	3	231	6	509
Total		2338		3555

variety of technical faults that the technicians had to cope with was higher. Also, a somewhat higher proportion of visits were call-outs (between 25 and 30 per cent) so demand was slightly less predictable. The real difference between domestic and business customers, according to Mark, was the nature of the contact between KTS technicians and customers. '*Business customers want to be involved in knowing the best way to use their systems. They want advice, and they want to know what you are doing. So, for example, if you install an update to the system control software, they usually want to be informed. They also either keep a servicing log themselves, or ask us to report on measures such as boiler efficiency, time between repairs, downtime due to failure or servicing (particularly important), and so on. Call-out response time is particularly important for them, but because there is usually someone always on their premises, it is easier to arrange a time to call for regular servicing.*'

Both Ros and Mark were disappointed that the design and installation business had been slow to take off. The one system designer that they had hired was proving an asset, and two of their technicians from the business servicing side of the operation had been moved over to installation work and were proving successful. '*It's a tight team of three at the moment*', said Mark, '*and that should give us enough capacity for the remainder of the year. But we will eventually need to recruit more technicians as business (hopefully) builds up.*' The extra workshop space that the firm had rented (on the same site) and some new equipment had allowed the design and installation team to adapt and customise boiler and control systems to suit individual customers' requirements. '*Many installers are owned by boiler manufacturers and can be guilty of pushing a standard solution on customers. With us, every system is customised to each customer's needs.*' (Mark Trenton)

KTS organisation

A small administrative office of four people reported directly to Ros and Mark, and helped manage accounting, HR, invoicing, contract maintenance and purchasing activities. The office was adjacent to a workshop space shared by the domestic and business boiler technicians. KTS employed 42 technicians in total. Nominally 26 of these worked on domestic boiler servicing and repair, and 16 on business boiler servicing and repair, yet there was some flexibility between the two groups. '*We are lucky that our technicians are usually reasonable about helping each other out*', said Mark. '*It is generally easier for the technicians used to serving business customers to serve domestic ones. They are not always as efficient as those used to domestic customers, but their customer-facing skills are usually better. Domestic boiler technicians do not always appreciate that business customers want more reassurance and information generally. Also, it is important for business customers to receive a full technical report within a couple of days of a visit. Domestic technicians are not used to doing that.*'

Improving service efficiency

Although both Ros and Mark were broadly happy with the way the business was developing, Mark in particular felt that they could be more efficient in how they organised themselves. '*Our costs have been increasing more or less in line with revenue growth, but we should really be starting to get some economies of scale. We need to improve our productivity, and I think we can achieve this by reducing waste. For example, we have found that our technicians can waste up to 30 per cent of their time on non-value-adding activities, such as form-filling or retrieving technical information.*'

Mark's solution was to tackle waste in a number of ways:

- Establish key performance indicators (KPIs) and simple metrics – performance measures must be clearly explained so that technicians understand the objectives that underlie their targets in terms of availability, utilisation and efficiency.
- Better forecasting – demand was forecast only in the simplest terms. Historical data to account for seasonality had not been used, nor had obvious factors, such as weather, been monitored.
- Slicker processes – administrative and other processes had been developed 'organically' with little consideration of efficiency.
- Better dispatching – dispatching (the allocation of jobs to individual technicians) was usually done on a simple 'first come, first served' basis without taking the efficient use of technicians' time into account. It was believed that both travel time and 'time to uptime' could be improved by better allocation of jobs.
- Better training – in the previous two years, three technicians had retired, one had been dismissed and two left for other jobs. Mark had found difficulty in replacing them with experienced people. It had become clear that it would become more important to hire inexperienced people and train them. In Mark's words, '*to get smart people with the right attitude and problem-solving skills, who don't mind getting their hands dirty, and give them the technical skills*'.

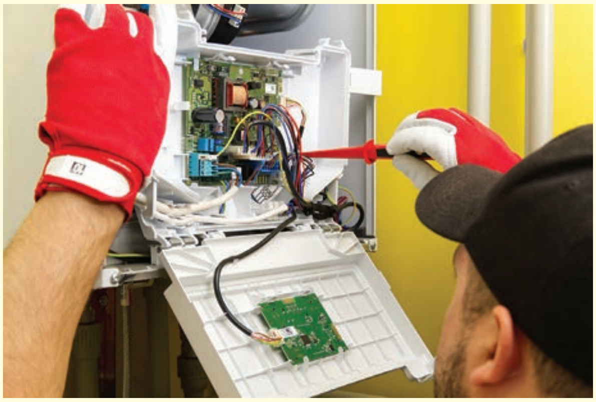
In addition to thinking about how best to improve efficiency, future market growth was also a concern. Two developments were occupying Ros and Mark's thoughts, one in the short to medium term, the other in the longer term.

Future growth – short to medium term

Demand had been growing steadily, largely by KTS winning business from smaller competitors. But Mark wondered whether the nature of what customers would want was changing. An opportunity had been suggested by one of KTS's oldest business customers. They had been approached by another HVAC company who had asked if they would be interested in a 'total' service where the company would both supply and operate a new heating system. In effect they were asking if KTS's customer would totally outsource their heating to them. It was an idea that Mark was intrigued by. '*I have heard about this type of deal before, but mainly for large businesses and offered by facilities management companies. It can involve companies like ours actually buying the heating system, installing it and taking responsibility for managing, not just the system itself, but actually how much energy is used. Exactly how it might work will, I guess, depend on the terms of the contract. Does the customer pay an amount per unit of energy used (perhaps linked to the wholesale price of energy)? Or does the customer simply pay a fixed amount for agreed operating characteristics, such as maintaining a particular temperature range? We would have to think carefully about the implications for us before offering such a service. The customer who told us about the approach does not want to desert us, but who knows what they might do in the future?*'

The future – longer term

According to the Committee on Climate Change (CCC), an independent advisory body that assisted the UK government in reaching required carbon levels, meeting the UK's target to reduce emissions would require reducing domestic emissions by at least 3 per cent per year – a challenging target. This would mean that within a few years it could become illegal to install gas boilers in new-build homes. One possible future that was discussed in the industry was a general move towards a hydrogen network (burning hydrogen produces no emissions and creates only water vapour and heat). However, a more likely future would probably involve combining different renewable technologies to provide low-carbon heat. The lowest-cost long-term solution could be to replace gas and oil boilers with hydrogen alternatives alongside electric heating generated from renewable sources such as air source or ground source heat pumps, which use small amounts of electricity to draw natural heat from either the air or ground. To make heat pumps effective, all existing and new-build homes would need to be made energy efficient by using far better levels of insulation.



VidEst/Shutterstock

Ros thought that these developments could prove far more challenging for KTS. '*Both Mark and I had assumed that we would be in this business for at least another 20 to 30 years. We both have families, so the long-term future of the business is obviously important to us. New heating technologies and fuels pose both opportunities and threats (yes, I've done an MBA) for us. Reducing fossil fuel consumption will definitely mean that we have to change what we do. And some aspects of demand may reduce. For example, ground source systems require little maintenance. But if there is going to be an upswing in the installation market, we need to be on top of it.*'

Questions

1. How would you position each of KTS's services on the 'four V' dimensions of volume, variety, variation and visibility?
2. What aspects of performance are important for KTS to win more servicing business?
3. How would you evaluate the potential of offering a 'total' service like the one the KTS customer had been offered?
4. What should KTS be doing to prepare for possible longer-term changes in the industry?

Applying the principles

Some of these exercises can be answered by reading the chapter. Others will require some general knowledge of business activity and some might require an element of investigation. **Model answers for the first two questions are available on the student companion website.** *For model answers to the other questions in this section, please ask your tutor/lecturer.*

1. Quentin Cakes make about 20,000 cakes per year in two sizes, both based on the same recipe. Sales peak at Christmas time when demand is about 50 per cent higher than in the more quiet summer period. Their customers (the stores that stock their products) order their cakes in advance through a simple internet-based ordering system. Knowing that they have some surplus capacity, one of their customers has approached them with two potential new orders.

The *Custom Cake* option – this would involve making cakes in different sizes where consumers could specify a message or greeting to be 'iced' on top of the cake. The consumer would give the inscription to the store who would email it through to the factory. The customer thought that demand would be around 1,000 cakes per year, mostly at celebration times such as Valentine's Day and Christmas.

The *Individual Cake* option – this option involves Quentin Cakes introducing a new line of about 10–15 types of very small cakes intended for individual consumption. Demand for this individual-sized cake was forecast to be around 4,000 per year, with demand likely to be more evenly distributed throughout the year than their existing products.

The total revenue from both options is likely to be roughly the same and the company has capacity to adopt only one of the ideas. But which one should it be?

2. Described as having '*revolutionised the concept of sandwich making and eating*', Pret A Manger opened their first shop in the mid-1980s, in London. Now they have over 130 shops in the UK, New York, Hong Kong and Tokyo. They say that their secret is to focus continually on quality, in all their activities. '*Many food retailers focus on extending the shelf life of their food, but that's of no interest to us. We maintain our edge by selling food that simply can't be beaten for freshness. At the end of the day, we give whatever we haven't sold to charity to help feed those who would otherwise go hungry.*' The first Pret A Manger shop had its own kitchen where fresh ingredients were delivered first thing every morning, and food was prepared throughout the day. Every Pret shop since has followed this model. The team members serving on the tills at lunchtime will have been making sandwiches in the kitchen that morning. They rejected the idea of a huge centralised sandwich factory, even though it could significantly reduce costs. Pret also own and manage all their shops directly so that they can ensure consistently high standards. '*We are determined never to forget that our hardworking people make all the difference. They are our heart and soul. When they care, our business is sound. If they cease to care, our business goes down the drain. We work hard at building great teams. We take our reward schemes and career opportunities very seriously. We don't work nights (generally), we wear jeans, we party!*'

a) Do you think Pret A Manger fully understand the importance of their operations management?

b) What evidence is there for this?

c) What kinds of operations management activities at Pret A Manger might come under the four headings of direct, design, deliver and develop?

3. Focus plastics originally made precision plastic components for the Aerospace sector, together with some basic (cheap) 'homeware' items such as buckets and dustpans. However, competition became intense in this market, so they decided to specialise in homewares; not just the basic products it had made previously, but a higher-quality, attractively designed and fashionable range. Its original basic range was sold under the 'Focus' brand and the new 'upmarket' range under the 'Concept' brand. The main customers for the Focus products were large homeware and 'do-it-yourself' retailers. The company delivered these products, in bulk, to their customers' fulfilment centres. By contrast, the Concept range was distributed to a newly developed network of exclusive stores, and kitchen equipment and speciality retailers. Within a year of launching the new Concept range, the company had over 3,000 retail outlets across Northern Europe with full point-of-sale display facilities and supported by press coverage. Within two years, Concept products were providing over 75 per cent of their revenue and 90 per cent of their profits. The Focus (basic) range continued to be produced, but as a drastically reduced range. The success of the Concept range enabled the company to replace their small injection-moulding machines with new larger ones that allowed them to use large multi-cavity moulds. This increased productivity by giving the capability to produce several products, or components, each machine cycle. Although cumbersome and more difficult to change over, the moulds were very efficient and gave a very high-quality product. With the same labour, they could make three items per minute on the old machines, and 18 items per minute on the modern ones using multi moulds. They also achieved high dimensional accuracy, excellent surface finish and extreme consistency of colour. Also, by standardising on single large machines, any mould could fit any machine. This was an ideal situation from a planning perspective, as they often had to make small quantities of Concept products at short notice.

Notwithstanding the success of the Concept range, the company had problems with its scheduling process. Its operations controller explained: '*We can change colours in 15 minutes, but because our moulds are large and technically complex, mould changes can take up to three hours. Good scheduling is important to maintain high plant utilisation. With a higher variety of complex products, batch sizes have reduced and it has brought down average utilisation. Often, we can't stick to schedules. Short-term changes are inevitable in a fashion market. Certainly better forecasts would help. . . but even our own promotions are sometimes organised at such short notice that we often get caught with stockouts. Yet, our finished stock levels are also growing fast*.'

Why might the company's stock of finished products be increasing while they are also suffering stockouts?

4. Focus plastics (see the previous question) moved into what it called 'design house partnerships' – design collaboration between their internal designers and Italian design houses, creative product designers who rarely manufacture or distribute their own products, relying on outsourcing to subcontractors. Using the design house's brand, they collaborated over the design of products, and then manufactured and distributed them. From the customer's point of view, the distribution arrangements appeared to belong to the design house itself. In fact, they were based exclusively on Focus' own call centre, warehouse and distribution resources. Using the four Vs model, how does the design house partnership offering contrast with the company's Focus and Concept offerings?

5. Find a copy of a financial newspaper (e.g. *Financial Times*, *Wall Street Journal*, *Economist*) and identify one company that is described in the paper that day. What do you think would be the main operations issues for that company?

Notes on chapter

1 The information for this example came from discussions with Tom Dyson and Olly Willans of Torchbox, to whom we are grateful for their advice and assistance.

2 The information on which this example is based was taken from the hotel's website, https://www.marinabaysands.com/ [accessed 14 September 2020].

3 Vargo, S.L. and Lusch, R.F. (2008) 'Service-dominant logic: continuing the evolution', *Journal of the Academy of Marketing Science*, 36 (1), pp. 1–10.

4 The information for this example is taken from: The Ellen Macarthur Foundation blog, https://www.ellenmacarthurfoundation.org/case-studies/selling-light-as-a-service [accessed 8 September 2020]; Phipps, L. (2018) 'How Philips became a pioneer of circularity-as-a-service', GreenBiz, 22 August; and press release, https://www.philips.com/a-w/about/news/archive/standard/news/press/2015/20150416-Philips-provides-Light-as-a-Service-to-Schiphol-Airport.html [accessed 8 September 2020].

5 In 2018 Philips Lighting changed its name to Signify, although it still uses the Philips brand for many of its products.

6 The phrase 'the triple bottom line' was first used in 1994 by John Elkington, the founder of a British consultancy called SustainAbility. Read Elkington, J. (1997) *Cannibals with Forks: The triple bottom line of 21st century business*, Capstone. Also good is Savitz, A.W. and Weber, K. (2006) *The Triple Bottom Line: How today's best-run companies are achieving economic, social and environmental success – and how you can too*, Jossey-Bass.

7 Jensen, M.C. (2001) 'Value maximization, stakeholder theory, and the corporate objective function', *Journal of Applied Corporate Finance*, 14 (3), pp. 8–21.

8 The information on which this example is based was taken from: the MSF website, http://www.msf.org.uk [accessed 8 September 2020]; and Beaumont, P. (2011) 'Médecins sans Frontières book reveals aid agencies' ugly compromises', *Guardian*, 20 November.

9 The information on which this example is based is taken from: Hernández, A. (2020) 'Learning from Adidas' Speedfactory blunder', Supplychaindive, 4 February, https://www.supplychaindive.com/news/adidas-speedfactory-blunder-distributed-operations/571678/ [accessed 8 September 2020]; Bain, M. (2019) 'Change of plan', Quartz, 11 November, https://qz.com/1746152/adidas-is-shutting-down-its-speedfactories-in-germany-and-the-us/ [accessed 8 September 2020].

10 Boudreau, J., Hopp, W., McClain, J.O. and Thomas, L.J. (2003) 'On the interface between operations and human resources management', *Manufacturing & Service Operations Management*, 5 (3), pp. 179–202.

Taking it further

Anupindi, R., Chopra, S., Deshmukh, S.D., Van Mieghem, J.A. and Zemel, E. ***(2013) Managing Business Process Flows, 3rd edition, Pearson.*** *Takes a 'process' view of operations, it's mathematical but rewarding.*

Barnes, D. ***(2018) Operations Management: An international perspective, Macmillan Education UK.*** *A text that is similar in outlook to this one, but with more of a (useful) international perspective.*

Hall, J.M. and Johnson, M.E. ***(2009) 'When should a process be art, not science?', Harvard Business Review, March.*** *One of the few articles that looks at the boundaries of conventional process theory.*

Holweg, M., Davies, J., De Meyer, A., Lawson, B. and Schmenner, R. ***(2018) Process Theory: The principles of operations management, Oxford University Press.*** *As the title implies, this is a book about theory. It is unapologetically academic, but does contain some useful ideas.*

Johnston, R., Shulver, M., Slack, N. and Clark, G. ***(2020) Service Operations Management, 5th edition, Pearson.*** *What can we say! A great treatment of service operations from the same stable as this text.*

Lewis, M.A. and Brown, A.D. ***(2012) 'How different is professional service operations management?', Journal of Operations Management, 30 (1–2), pp. 1–11.*** *It's an academic article, but it does say something interesting about professional service operations.*

Slack, N. ***(2017) The Operations Advantage: A practical guide to making operations work, Kogan Page.*** *Similar coverage to this text (and just as brilliant), but aimed very much at a practitioner audience.*

2 Operations and strategic impact

Introduction

In the long term, the major (and some would say, only) objective for operations and processes is to provide a business with some form of strategic advantage. That is the reason why the management of a business's processes and operations and its intended overall strategy must be logically connected. Yet for many in business, the very idea of an 'operations strategy' is a contradiction in terms. After all, to be involved in the strategy process is the complete opposite of those day-to-day tasks and activities associated with being an operations manager. Nevertheless, it is also clear that operations can have a real strategic impact. Many enduringly remarkable enterprises, from Apple to Zara, use their operations resources to gain long-term strategic success. Such firms have found that it is the way they manage their operations that sets them apart from, and above, their competitors. But if operations strategy can be the making of business success, a failure to have one, or a failure to adapt it to circumstances, can break a business. Without a strong link with overall strategy, operations and processes will be without a coherent direction. And without direction a business may finish up making internal decisions that either do not reflect strategy, or that conflict with each other, or both. So, although operations and process management is largely 'operational', it also has a strategic dimension that is vital if operations is to fulfil its potential to contribute to competitiveness. Figure 2.1 shows the position of the ideas described in this chapter in the general model of operations management.

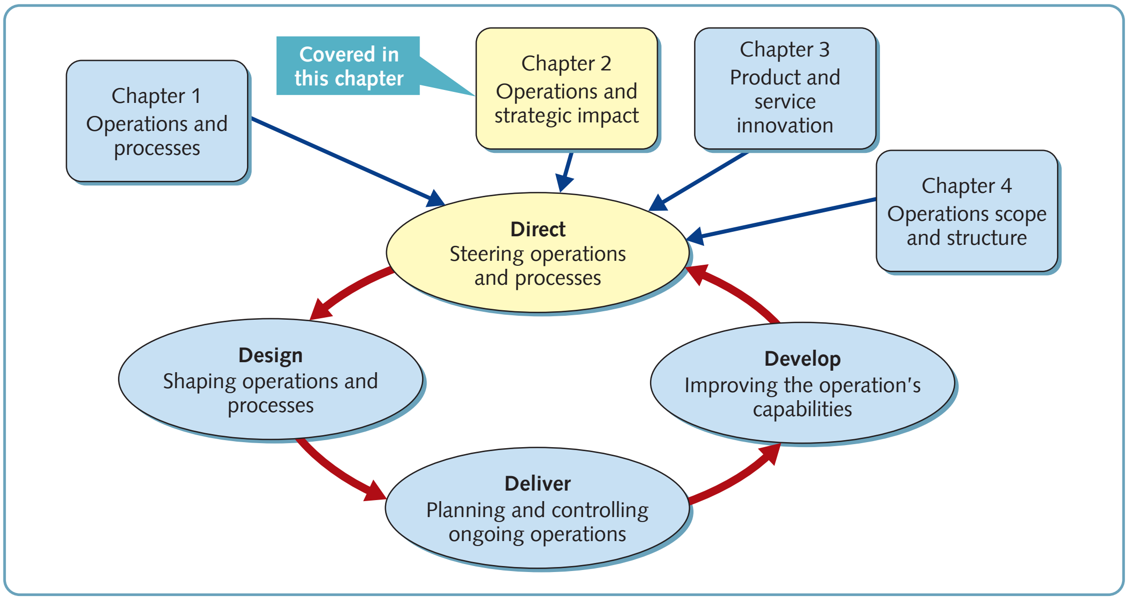
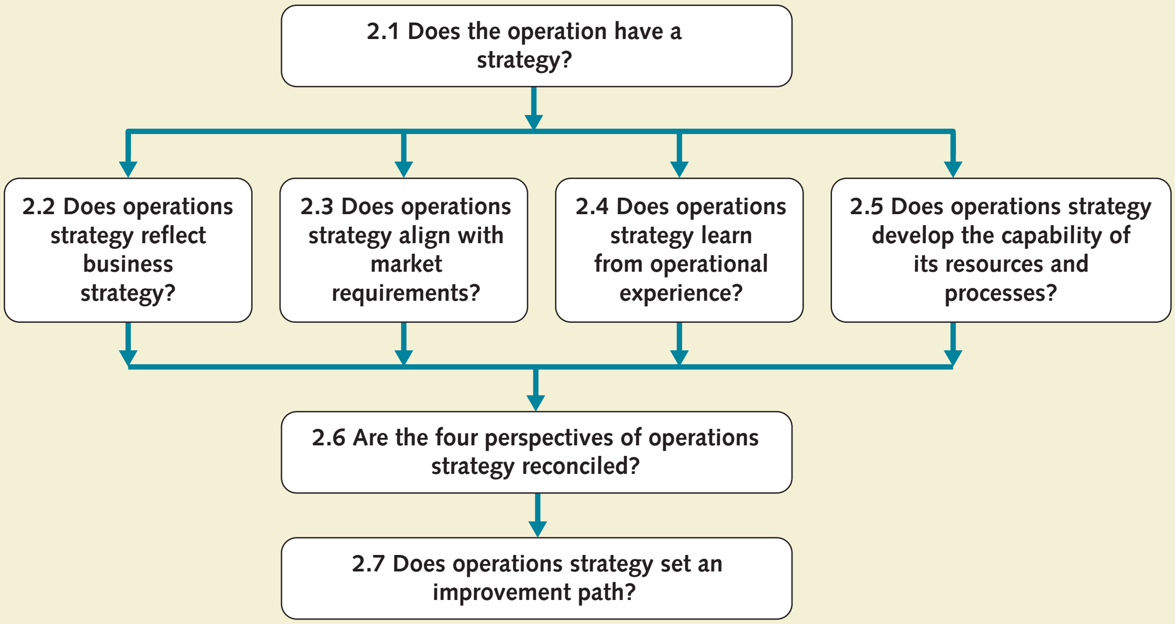


Figure 2.1 Operations strategy is the pattern of decisions and actions that shapes the long-term vision, objectives and capabilities of the operation and its contribution to overall strategy

EXECUTIVE SUMMARY



2.1 Does the operation have a strategy?

Operations strategy is the pattern of decisions and actions that shapes the long-term vision, objectives and capabilities of the operation and its contribution to overall strategy. It is the way in which operations resources are developed over the long term to create sustainable competitive advantage for the business. Increasingly, many businesses are seeing their operations strategy as one of the best ways to differentiate themselves from competitors. Even in those companies that are marketing-led (such as fast-moving consumer goods), an effective operations strategy can add value by allowing the exploitation of market positioning. Strategies are always difficult to identify because they have no presence in themselves, but are identified by the pattern of decisions that they generate. Nevertheless, one can identify what an operations strategy should do. First, it should take significant stakeholders into account. They are the people and groups who have a legitimate interest in the operation's strategy. They include employees, customers, society or community groups, shareholders, suppliers and industry regulators. Second, it should articulate a vision for the operations contribution. This is a clear statement of how operations intend to contribute value for the business. A common approach to summarising operations contribution is the Hayes and Wheelwright Four-Stage Model (see Figure 2.2).

2.2 Does operations strategy reflect business strategy (top-down)?

A top-down perspective often identifies three related levels of strategy: corporate strategy that should position the corporation in its global, economic, political and social environment; business strategy that sets out each business's individual mission and objectives; and functional strategy that considers what part each function should play in contributing to the strategic objectives of the business. Any top-down perspective on operations strategy should attempt to achieve both correspondence (it is consistent with business strategy) and coherence (its various elements all 'pull' in the same direction). One way of achieving this is to devise a 'business model' that integrates with an 'operating model'.

2.3 Does operations strategy align with market requirements (outside-in)?

Operations strategy should reflect the market position of the business. And because companies compete in different ways, the operations function should respond by providing the ability to perform in a manner that is appropriate for the intended market position. This is called a market (or outside-in) perspective on operations strategy. It involves translating market requirements into 'performance objectives' (quality, speed, dependability, flexibility and cost), the exact definition of which will be different for different operations. These may be refined by distinguishing between those performance objectives that are 'order-winners' (those things that directly and significantly contribute to winning business) and those that are 'qualifiers' (the aspects of competitiveness where the operation's performance has to be above a particular level just to be considered by the customer).

2.4 Does operations strategy learn from operational experience (bottom-up)?

Operations may adopt a particular strategic direction not because of any formal high-level decision-making, but because the ongoing experience of providing products and services to customers at an operational level convinces them that it is the right thing to do. This is the 'bottom-up' perspective of operations strategy. It stresses the concept of 'emergent strategies' where strategy-making can be relatively unstructured and fragmented. The important attributes for shaping strategy from the bottom up are an ability to capture the learning that should come from routine operations activities and being able to transform that learning into strategically valuable knowledge.

2.5 Does operations strategy develop the capability of its resources and processes (inside-out)?

In the longer term, operations strategy must build the capabilities within its resources and processes that will allow the business to provide something to the market that its competitors find difficult to imitate or match. The idea of the basis of long-term competitive capabilities deriving from the operation's resources and processes is called the resource, or 'inside-out', perspective on operations strategy. It is very much influenced by the resource-based view (RBV) of the firm. 'Difficult-to-imitate' resources can be classified as 'strategic' if they are scarce, or imperfectly mobile, or imperfectly imitable, or imperfectly substitutable.

2.6 Are the four perspectives of operations strategy reconciled?

None of the four perspectives individually can give a full picture of any operations strategy. But together they do provide a good idea of how operations are contributing strategically. Yet, the four perspectives must be reconciled. One way to do this is to construct an operations strategy matrix. This is a simple model that describes operations strategy as the intersection of a company's performance objectives and the strategic decisions that it makes. Alternatively, the 'line of fit' model (see Figure 2.11) can be used to map the balance between requirements of the market and the capabilities of the operation over time. Ideally, there should be a reasonable degree of alignment, or 'fit', between the two.

2.7 Does operations strategy set an improvement path?

The purpose of operations strategy is to improve the business's performance relative to its competitors' in the long term. It therefore must provide an indication of how this improvement is to take place. This is best addressed by considering the trade-offs between performance objectives in terms of the 'efficient frontier' model. This describes operations strategy as a combination of repositioning performance along an existing efficient frontier, and increasing overall operations effectiveness by overcoming trade-offs to expand the efficient frontier.

2.1 Diagnostic question: *Does the operation have a strategy?*

Let us start with the meaning of 'strategy'. Surprisingly, it is not particularly easy to define. Linguistically the word derives from the Greek word '*strategos*' meaning 'leading an army'. And although there is no direct historical link between Greek military practice and modern ideas of strategy, the military metaphor is powerful. Both military and business strategy can be described in similar ways, and include some of the following:

- Setting broad objectives that direct an enterprise towards its overall goal.
- Planning the path (in general rather than specific terms) that will achieve these goals.
- Stressing long-term rather than short-term objectives.
- Dealing with the total picture rather than stressing individual activities.
- Being detached from, and above, the confusion and distractions of day-to-day activities.

All these points apply to operations strategy. It is the 'pattern of decisions and actions that shape the long-term vision, objectives and capabilities of the operation and its contribution to the overall strategy of the business'. Yet, the term 'operations strategy' sounds at first like a contradiction. How can 'operations', a subject that is generally concerned with the day-to-day creation and delivery of goods and services, be strategic? 'Strategy' is usually regarded as the opposite of those day-to-day routine activities. But, as we indicated previously, 'operations' is not the same as 'operational'. Operations are the resources and processes that create products and services. It is 'operational' that is the opposite of strategic. It means day-to-day, detailed and often localised.

Perhaps more significantly, many of the businesses that seem to be especially successful, and that appear to be sustaining their success into the longer term, have a clear and often inventive operations strategy. Just look at some of the high-profile companies quoted in this text, or that feature in the business press. It is not just that their operations strategy provides these companies with adequate support; it is their operations strategy that is the pivotal reason for their competitive superiority. Just as revealing, when companies stumble it is often because they have either taken their eye off the operations ball, or failed to appreciate its importance in the first place. Operations strategy can make or break a business. More generally, all enterprises, and all parts of the enterprise, need to prevent strategic decisions being frustrated by poor operational implementation. And this idea leads us to the second purpose of this chapter (and indeed the text as a whole): to show how, by using the principles of operations strategy, all parts of any business (and all functions of a business) can contribute effectively to the overall success of that business. So the idea of 'operations strategy' has two different but related

Case example

Surrey Satellite Technology Limited (SSTL)[1]

Sometimes a firm's operations strategy can change an industry. Look at space satellites. They can be expensive – very expensive. And in the early days of space missions, only superpowers could afford to develop and launch them. The conventional wisdom was that space was such a hostile environment that satellites would have to be constructed using only expensive, specially developed components. Yet in the late 1970s this assumption was challenged by Sir Martin Sweeting, who then was studying for his PhD at the University of Surrey in the UK. The aerospace research team in the Electrical Engineering Department at the University of Surrey had built its first satellite using commercial off-the-shelf components. It was about as big as two microwave ovens, weighing in at 72 kg (by contrast, some of the satellites being launched by government space agencies were as large as a London double-decker bus). It was launched in 1981 with the help of NASA. The team followed this up in 1984 with a second satellite that they built in just six months. A year later Surrey Satellite Technology Limited (SSTL) was formed. The firm's vision was to open up the market for space exploration by pioneering the use of small and relatively cheap, but reliable, satellites built from readily available off-the-shelf components. It was a revolutionary idea. Now SSTL is the world's leading small satellite company, and has delivered space missions for a whole range of applications including Earth observation, science, communications and in-orbit technology demonstration. The company is at the forefront of space innovation, exploiting advances in technologies and challenging conventions to bring affordable space exploration to international customers. The company, which employs more than 500 staff, has launched over 40 satellites and is based across four sites in south-east England. Since 2014 SSTL has been an independent company within the Airbus Defence and Space group.

How has SSTL achieved this success from such small beginnings? Partly because it was an early player in the market, having the vision to see that there would be a market for small satellites that could serve the ambitions of smaller countries, companies, research groups and even schools. But in addition, it has always been innovative in finding ways of keeping the cost of building the satellites down. SSTL pioneered the low-cost, low-risk approach to delivering operational satellite missions within short development timescales and with the capability that potential customers wanted. Particularly important was the company's use of commercial off-the-shelf technology. In effect, using industry-standard parts meant exploiting the (often enormous) investments by consumer-electronics companies, auto-part manufacturers and others who had developed complex components for their products. Even if this sometimes limited what a satellite could do, it provided the scale economies that would be impossible if they were designing and making customised components from scratch. *'We were being parasitic, if you like,'* admits Sir Martin.

However, not all commercially available components made for terrestrial use are up to coping with conditions in space, which is a hugely important issue. Reliability is essential in a satellite. (It's difficult to repair them once in space.) And even though off-the-shelf components and systems have become increasingly reliable, they must be rigorously tested to make sure that they are up to the severe conditions found in space. Knowing which bits can be used and which cannot is an important piece of knowledge. Yet, although individual components and systems are often bought off the shelf, the company does most of its operations activities itself. This allows SSTL to provide a complete in-house design, manufacture, launch and operation service, as well as a range of advice, analysis and consultancy services. *'What distinguishes us is our vertically integrated capability, from design and research to manufacturing and operations,'* says Sir Martin. *'We don't have to rely on suppliers, although of course we buy-in components when that is advantageous.'* And innovation? It's still as important as it was at the company's start.

Paul Fleet/Alamy Stock Photo

meanings. The first is concerned with the operations function itself, and how it can contribute to strategic success. The second is concerned with how any function can develop its resources and processes to establish its strategic role.

Look at the case example on SSTL. They exhibited several important characteristics that one finds in many successful companies. First, they were innovative in how they exploited the potential of emerging technologies. They realised that, with clever design, relatively inexpensive components could achieve a large part of what far more costly components could. Second, SSTL linked the development of their operations to a well-defined idea of what their customers wanted. All aspects of their operations strategy had clear customer benefits. Third, they learned from their experience. A key part of SSTL's core knowledge is which components can stand up to conditions in space. Finally, and most significantly, they actually did have an operations strategy; they realised the importance of strategically directing their operations resources. SSTL did not ignore the strategic potential of their operations resources and processes.

How do you judge an operations strategy?

There are some problems in asking this apparently simple question. In most operations management decisions you can see what you are dealing with. You can touch inventory, talk to people, programme machines, and so on. But strategy is different. You cannot see a strategy, feel it or touch it. Also, whereas the effects of most operations management decisions become evident relatively fast, it may be years before an operations strategy decision can be judged to be a success or not. Moreover, any 'strategy' is always more than a single decision. Operations strategy will be revealed in the total pattern of decisions that a business takes in developing its operations in the long term. Nevertheless, the question is an obvious starting point and one that must be addressed by all operations.

An operations strategy should take significant stakeholders into account

All operations have stakeholders. They are the people and groups who have a legitimate interest in the operation's strategy. Some are internal (employees); others are external (customers, society or community groups, and a company's shareholders). External stakeholders may have a direct commercial relationship with the organisation (suppliers and customers); others may not (industry regulators). In not-for-profit operations, these stakeholder groups can overlap. So, voluntary workers in a charity may be employees, shareholders and customers all at once. However, in any kind of organisation, it is a responsibility of the operations function to understand the (often conflicting) objectives of its stakeholders and set its objectives accordingly. Yet, although all stakeholder groups, to different extents, will be interested in operations performance, they are likely to have very different views on which aspect of performance is important. Table 2.1 identifies typical stakeholder requirements. But stakeholder relationships are not just one-way. It is also useful to consider what an individual organisation or business wants of the stakeholder groups themselves. Some of these requirements are illustrated in Table 2.1. The dilemma with using this wide range of stakeholders to judge performance is that organisations, particularly commercial companies, have to cope with the conflicting pressures of maximising profitability on the one hand with the expectation that they will manage in the interests of (all or part of) society in general with accountability and transparency. Even if a business wants to reflect aspects of performance beyond its own immediate interests, how is it to do it?

OPERATIONS PRINCIPLE
Operations strategy should take the requirements of significant stakeholders into account.

Corporate social responsibility (CSR)

Strongly related to the stakeholder perspective of operations performance is that of corporate social responsibility (generally known as CSR). It is generally taken to mean listening and responding to the needs of a company's stakeholders, including the requirements of sustainable

Table 2.1 Typical stakeholders' performance objectives

Stakeholder groups	*What stakeholders want from the operation*	*What the operation wants from stakeholders*
Shareholders	Return on investment Stability of earnings Liquidity of investment	Investment capital Long-term commitment
Directors/top management	Low/acceptable operating costs Secure revenue Well-targeted investment Low risk of failure Future innovation	Coherent, consistent, clear and achievable strategies Appropriate investment
Staff	Fair wages Good working conditions Safe work environment Personal and career development	Attendance Diligence/best efforts Honesty Engagement
Staff representative bodies (e.g. trade unions)	Conformance with national agreements Consultation	Understanding Fairness Assistance in problem solving
Suppliers (of materials, services, equipment, etc.)	Early notice of requirements Long-term orders Fair price On-time payment	Integrity of delivery, quality and volume Innovation Responsiveness Progressive price reductions
Regulators (e.g. financial regulators)	Conformance to regulations Feedback on effectiveness of regulations	Consistency of regulation Consistency of application of regulations Responsiveness to industry concerns
Government (local, national, regional)	Conformance to legal requirements Contribution to (local/national/regional) economy	Low/simple taxation Representation of local concerns Appropriate infrastructure
Lobby groups (e.g. environmental lobby groups)	Alignment of the organisation's activities with whatever the group are promoting	No unfair targeting Practical help in achieving aims (if the organisation wants to achieve them)
Society	Minimise negative effects from the operation (noise, traffic, etc.) and maximise positive effects (jobs, local sponsorship, etc.)	Support for organisation's plans

development, and building good relationships with employees, suppliers and wider society. This issue of how broader social performance objectives can be included in operations management's activities is of increasing importance, both from an ethical and a commercial point of view. However, converting the CSR concept into operational reality presents considerable difficulties, although several attempts have been made. For example, the 'triple bottom line' approach that we described in the previous chapter is one of the best-known attempts to integrate economic, environmental and social impacts.

Case example

Danone's path to B Corporation[2]

The food industry is not always the most popular with some consumers, who doubt the environmental sustainability of industrial-scale food and drink production. A common view among food bloggers and radical activists is that the giants of the food industry, with their large factories and long supply chains, have disconnected consumers from the 'natural' source of their nourishment. Perhaps surprisingly, this view is shared by some in the food industry, and some are actively trying to balance financial, environmental and social objectives. Take Danone for example, the French food company that employs over a hundred thousand people, and markets over €25 billion worth of its products, such as Activia yogurt and Evian mineral water, throughout 130 countries. It is one of Europe's largest and most well-known food companies, that built its reputation on a health-focused portfolio of food products. It is a world leader in several food-related businesses, including dairy and plant-based products, early-life nutrition, medical nutrition and waters. Yet, the company believes that it should be looking beyond simply maximising profitability alone. It rejects the idea that a firm exists mainly to maximise the return to its owners, the shareholders. Its commitment to following a broader set of corporate objectives is exemplified by its ambition to become one of the first certified 'B Corp' multinationals. Certified B Corporations are businesses that meet the highest standards of verified social and environmental performance, public transparency and legal accountability to balance profit and purpose. The B Corp movement was launched in the US in 2006, and has been gaining momentum around the world. There are more than 2,300 B Corps in 50 countries. B Corps use profits and growth as a means to a greater end: positive impact for their employees, communities and the environment. Danone decided to partner with B Lab to plan the most appropriate roadmap toward this goal. B Lab is a non-profit organisation that accredits B Corp certification to for-profit companies that demonstrate high standards of social and environmental performance. And Danone has actively translated these goals into action. It disposed of subsidiaries that produced biscuits, chocolate and beer, and is trying to become carbon neutral with its water brands. It has also invested in a way to make recycled plastic (which is often an unappealing grey colour) more attractive to consumers.

An operations strategy should articulate a vision for the operations contribution

The 'vision' for an operation is a clear statement of how operations intend to contribute value for the business. It is not a statement of what the operation wants to achieve (those are its objectives), but rather an idea of what it must become and what contribution it should make.

> **OPERATIONS PRINCIPLE**
> ***Operations strategy should articulate a 'vision' for the operations function's contribution to overall strategy.***

A common approach to summarising operations contribution is the Hayes and Wheelwright Four-Stage Model. The model traces the progression of the operations function from what is the largely negative role of 'Stage 1' operations to it becoming the central element of competitive strategy in excellent 'Stage 4' operations. Figure 2.2 illustrates the four steps involved in moving from Stage 1 to Stage 4.

Stage 1: Internal neutrality

This is the very poorest level of contribution by the operations function. The other functions regard it as holding them back from competing effectively. The operations function is inward-looking and at best reactive, with very few positives to contribute towards competitive success. Its goal is to be ignored. At least then it isn't holding the company back in any way. Certainly, the rest of the organisation would not look to operations as the source of any originality, flair or competitive drive. Its vision is to be 'internally neutral', a position it attempts to achieve not by anything positive but by avoiding the bigger mistakes.

Stage 2: External neutrality

The first step of breaking out of Stage 1 is for the operations function to begin comparing itself with similar companies or organisations in the outside market. This may not immediately take

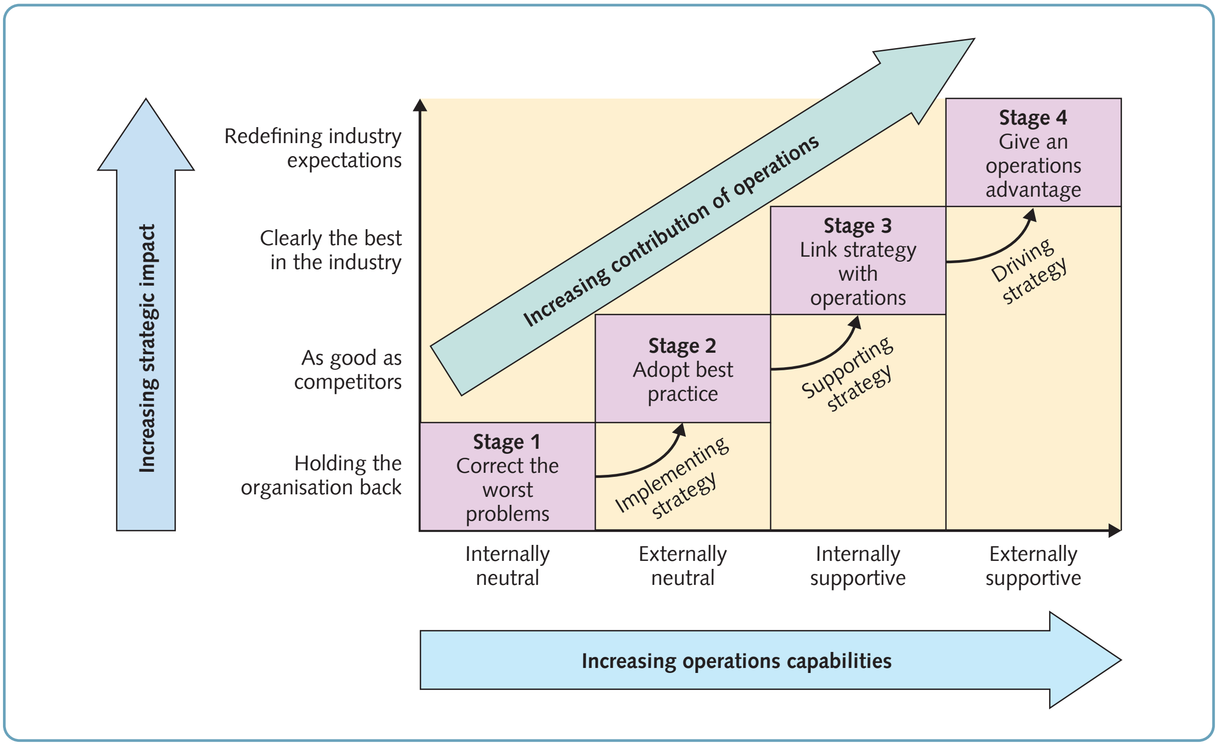


Figure 2.2 **Hayes and Wheelwright's Four-Stage Model of operations contribution sees operations as moving from implementation of strategy, through to supporting strategy and finally to driving strategy**

it to the 'first division' of companies in the market, but at least it is measuring itself against its competitors' performance and trying to be 'appropriate' by adopting 'best practice' from them. Its vision is to become 'up to speed' or 'externally neutral' with similar businesses in its industry by adopting 'best practice' ideas and norms of performance from others.

Stage 3: Internally supportive

Stage 3 operations have probably reached the 'first division' in their market. They may not be better than their competitors on every aspect of operations performance but they are broadly up with the best. Yet, the vision of Stage 3 operations is to be clearly and unambiguously the very best in the market. They may try to achieve this by gaining a clear view of the company's competitive or strategic goals and developing 'appropriate' operations resources to excel in the areas in which the company needs to compete effectively. The operation is trying to be 'internally supportive' by providing a credible operations strategy.

Stage 4: Externally supportive

Stage 3 used to be taken as the limit of the operations function's contribution. Yet, the model captures the growing importance of operations management by suggesting a further stage – Stage 4. The difference between Stages 3 and 4 is subtle, but important. A Stage 4 company is one where the vision for the operations function is to provide the foundation for competitive success. Operations look to the long term. It forecasts likely changes in markets and supply and, over time, it develops the operations-based capabilities that will be required to compete in future market conditions. The operations function is becoming central to strategy-making. Stage 4 operations are creative and proactive. They are innovative and capable of adaptation as markets change. Essentially, they are trying to be 'one step ahead' of competitors in the way that they create products and services and organise their operations – what the model terms being 'externally supportive'.

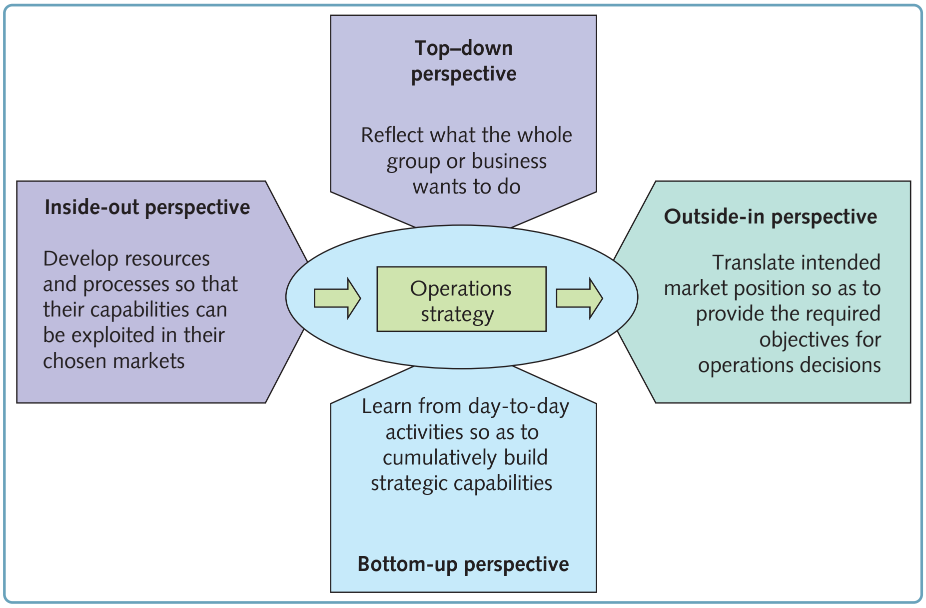


Figure 2.3 **The four perspectives on operations strategy**

The four perspectives on operations strategy

So, as an enterprise moves through each of these stages, what exactly should an operations strategy do? Different authors have slightly different views and definitions of operations strategy. However, between them, four 'perspectives' emerge (see Figure 2.3):

1. Operations strategy should reflect what the whole group or business wants to do in a 'top-down' manner.
2. Operations strategy should translate the enterprise's intended market position so as to provide the required objectives for operations decisions (sometimes called the 'outside-in' perspective).
3. Operations strategy should learn from day-to-day activities so as to cumulatively build strategic capabilities in a 'bottom-up' manner.
4. Operations strategy should develop its resources and processes so that their capabilities can be exploited in their chosen markets (sometimes called the 'inside-out' perspective).

None of these four perspectives alone gives the full picture of what operations strategy is. But together they provide some idea of the pressures that go to form the content of operations strategy. We will treat each of them in turn.

2.2 Diagnostic question: *Does operations strategy reflect business strategy (top-down)?*

A top-down perspective often identifies three related levels of strategy: corporate, business and functional. A corporate strategy should position the corporation in its global, economic, political and social environment. This will consist of decisions about what types of business the group wants to be involved in, what parts of the world it wants to operate in, how to allocate its cash between its various businesses, and so on. Each business unit within the corporate group will also need

> **OPERATIONS PRINCIPLE**
> ***Operations strategies should reflect top-down corporate and/or business objectives.***

to put together its own business strategy that sets out its individual mission and objectives. This business strategy guides the business in relation to its customers, markets and competitors, and also defines its role within the corporate group of which it is a part. Similarly, within the business, functional strategies need to consider what part each function should play in contributing to the strategic objectives of the business. The operations, marketing, product/service development and other functions will all need to consider how best they should organise themselves to support the business's objectives. This is why it is often called the 'top-down' perspective on operations strategy.

Although this rather neat relationship between the levels of corporate, business and operations strategy may seem a little 'theoretical', it is still a powerful idea. What it is saying is that in order to understand strategy at any level, one has to place it in the context of what it is trying to do (the level above) and how it is trying to do it (the level below). At any level, a good top-down perspective should provide clarity and connection. It should clarify what an operations strategy should be prioritising, and give some guidance on how the strategy is to be achieved.

Correspondence and coherence

However, developing any functional strategy from a business strategy is not a straightforward task. There are ambiguities to clarify and conflicts to be reconciled. Inevitably, business strategy consists of aggregated and approximate objectives. It should give an overall direction, but cannot spell out every detail of how a function should interpret its objectives. Yet, there should be a clear, explicit and logical connection between each functional strategy and the business strategy in which they operate. Moreover, there should also be a clear, explicit and logical connection between a functional strategy and the decisions taken within the function. In other words, there should be clear correspondence between a business's strategy and its operations strategy, as there should also be between an operations strategy and the individual decisions taken within the operations function.

But although correspondence between the levels of strategy is necessary, it is not all that is required. Operations strategy must also be coherent, both with other functional strategies and within itself. Coherence means that those choices made across or within functions should not pull it in different directions. All decisions should complement and reinforce each other in the promotion of the business's and the operations objectives. Figure 2.4 illustrates these two ideas of correspondence and coherence.

The concepts of the 'business model' and the 'operating model'

Two concepts have emerged over the last few years that are useful in understanding the top-down perspective on operations strategy (or at least the terms are new – one could argue that the ideas are far older). These are the concepts of the 'business model' and the 'operating model'.

Put simply, a 'business model' is the plan that is implemented by a company to generate revenue and make a profit (or fulfil its social objectives if a not-for-profit enterprise). It includes the various parts and organisational functions of the business, as well as the revenues it generates and the expenses it incurs. In other words, what a company does and how it makes money from doing it. More formally, it is 'a conceptual tool that contains a big set of elements and their relationships and allows [the expression of] the business logic of a specific firm. It is a description of the value a company offers to one or several segments of customers and of the architecture of the firm and its network of partners for creating, marketing, and delivering this value and relationship capital, to generate profitable and sustainable revenue streams.'

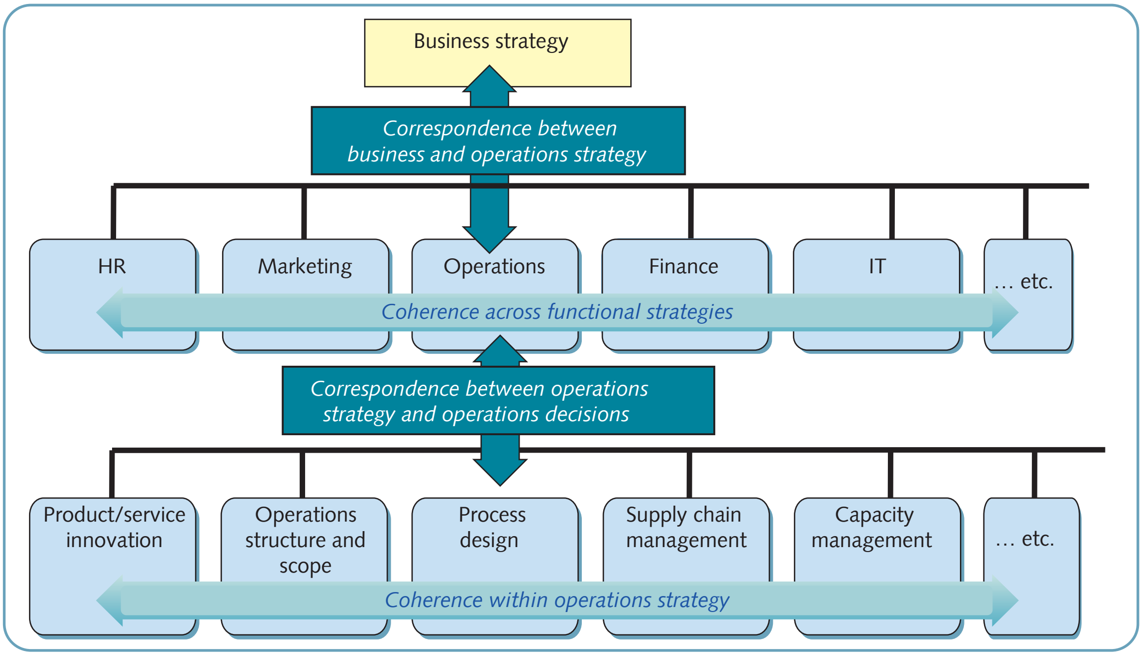


Figure 2.4 Correspondence and coherence are the two requirements of the top-down perspective of operations strategy

One synthesis of literature shows that business models have a number of common elements:

1. The value proposition of what is offered to the market.
2. The target customer segments addressed by the value proposition.
3. The communication and distribution channels to reach customers and offer the value proposition.

Worked example

Innovation at Micraytech (Part 1, top-down)

Micraytech is a metrology systems company that develops integrated measurement systems for large international clients in several industries; it is part of the Micray Group, which includes several high-tech companies. It has grown through a strategy of providing products with a high degree of technical excellence and innovation, together with an ability to customise its systems and offer technical advice to its clients. The Group has set ambitious growth targets for the company over the next five years and has relaxed its normal 'return on sales' targets to help it achieve this. As part of this strategy, Micraytech attempted to be the first in the market with all appropriate new technical innovations. From a top-down perspective, its operations function, therefore, needed to be capable of coping with the changes that constant product innovation would bring. It developed processes that were flexible enough to develop and assemble novel components and systems, while integrating them with software innovations. The company's operations staff realised that they needed to organise and train their staff to understand the way technology is developing so that they could put in place the necessary changes to the operation. It also needed to develop relationships with both existing and potentially new suppliers who could respond quickly when supplying new components. The top-down logic here is that everything about the operation – its processes, staff and its systems and procedures – must, in the short term, do nothing to inhibit, and in the long term actively develop, the company's competitive strategy of growth through innovation.

4. The relationships established with customers.
5. The core capabilities needed to make the business model possible.
6. The configuration of activities to implement the business model.
7. The partners and their motivations for coming together to make a business model happen.
8. The revenue streams generated by the business model constituting the revenue model.
9. The cost structure resulting from the business model.

One can see that this idea of the business model is broadly analogous to the idea of a 'business strategy', but implies more of an emphasis on how to achieve an intended strategy as well as exactly what that strategy should be.

The concept of an 'operating model' is more operational than that of a 'business model' and there is no universally agreed definition. Here we define it as a 'high-level design of the organisation that defines the structure and style which enables it to meet its business objectives'. Ideally, an operating model should provide a clear, 'big picture' description of what the organisation does and how it does it. It defines how the critical work of an organisation is carried out. It should provide a way to examine the business in terms of the key relationships between business functions, processes and structures that are required for the organisation to fulfil its mission. Unlike the concept of a business model, which often assumes a profit motive, the operating model philosophy can be applied to organisations of all types – including large corporations, not-for-profit organisations and the public sector.

Again, there is no universally agreed list of elements that an operating model should include and different organisations focus on different things, but many of the following elements are often included:

- Key performance indicators (KPIs) – with an indication of the relative importance of performance objectives.
- Core financial structure – P&L, new investments and cash flow.
- The nature of accountabilities for products, geographies, assets, etc.
- The structure of the organisation – sometimes expressed as capability areas rather than functional roles.
- Systems and technologies.
- Processes, responsibilities and interactions.
- Key knowledge and competence.

Note two important characteristics of an operating model. First, it does not respect conventional functional boundaries as such. In some ways the concept of the operating model reflects the idea that we proposed in Chapter 1. Namely, that all managers are operations managers and all functions can be considered as operations because they comprise processes that deliver some kind of service. An operating model is like an operations strategy, but applied across all functions and domains of the organisation. Second, there are clear overlaps between the 'business model' and the 'operating model', the main difference being that an operating model focuses more on how an overall business strategy is to be achieved. Also, operating models are rarely designed from first principles. Some kind of understood 'way of doing things' will already exist. This is why operating models often have an element of implied change or transformation of the organisation's resources and processes. Often the term 'target operating model' (TOM) is used to describe the way the organisation should operate in the future if it is going to achieve its objectives and make a success of its business model. Figure 2.5 illustrates the relationship between business and operating models.

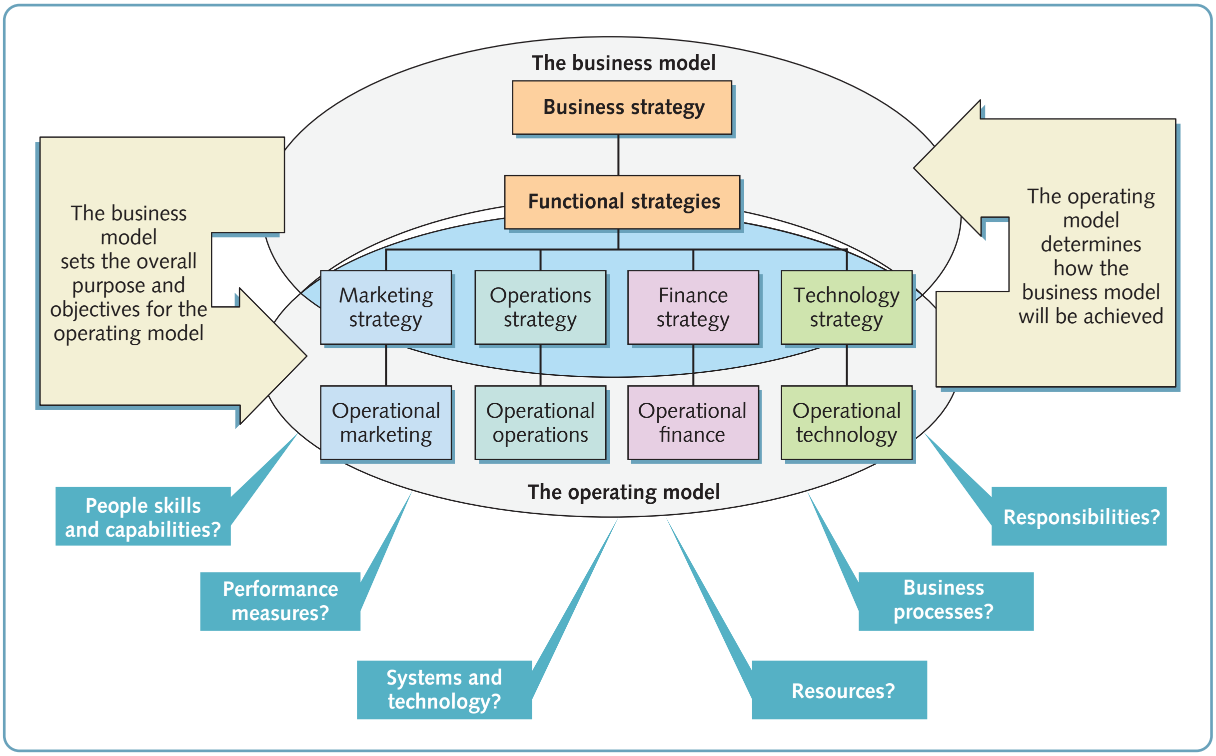


Figure 2.5 **The concepts of the 'business model' and the 'operating model' overlap – with the operating model indicating how processes, resources, technology, people, measures and responsibilities are to be organised to support the business model**

2.3 Diagnostic question: *Does operations strategy align with market requirements (outside-in)?*

Any operations strategy should reflect the intended market position of the business. Companies compete in different ways: some compete primarily on cost, others on the excellence of their products or services, others on high levels of customer service, and so on. The operations function should therefore respond to this by providing the ability to perform in a manner that is appropriate for the intended market position. This is a market (or outside-in) perspective on operations strategy.

An operations strategy should define operations performance objectives

Operations add value for customers and contribute to competitiveness by being able to satisfy the requirements of their customers. The most useful way to do this is to use the 'operational' performance measures that we discussed briefly in the previous chapter. All of these, to a greater or lesser extent, will affect customer satisfaction and define market positioning in terms that have meaning in an operations context. As a reminder, the five performance objectives are:

- *Quality* – producing error-free goods and services that are 'fit for their purpose'.
- *Speed* – minimising the time between a customer asking for goods and services and the customer receiving them in full.
- *Dependability* – keeping the delivery promises that have been made to customers.

- *Flexibility* – the ability to vary or adapt the operation's activities to cope with unexpected circumstances or to give customers individual treatment, or to introduce new products or services.
- *Cost* – producing goods and services at a cost that enables them to be priced appropriately for the market, while still allowing a return to the organisation (or, in a not-for-profit organisation, giving good value to the taxpayers or whoever is funding the operation).

The exact meaning of performance objectives is different in different operations

Different operations will have different views of what each of the performance objectives actually mean. Table 2.2 looks at how two operations, an insurance company and a steel plant, define each performance objective. For example, the insurance company sees quality as being at least as much about the manner in which their customers relate to their service, as it does about the absence of technical errors. The steel plant, on the other hand, while not ignoring quality of service, primarily emphasises product-related technical issues. Although they are selecting from the same pool of factors, which together constitute the generic performance objective, they will emphasise different elements.

OPERATIONS PRINCIPLE
Operations performance objectives can be grouped together under quality, speed, dependability, flexibility and cost.

Sometimes operations may choose to re-bundle elements using slightly different headings. For example, it is not uncommon in some service operations

Table 2.2 Aspects of performance objectives for two different kinds of operations

Insurance company *Aspects of each performance objective include:*	*Performance objectives*	*Steel plant* *Aspects of each performance objective include:*
• Professionalism of staff • Friendliness of staff • Accuracy of information • Ability to change details in future	**Quality**	• Percentage of products conforming to their specification • Absolute specification of products • Usefulness of technical advice
• Time for call centre to respond • Prompt advice response • Fast quotation decisions • Fast response to claims	**Speed**	• Lead time from enquiry to quotation • Lead time from order to delivery • Lead time for technical advice
• Reliability of original promise date • Customers kept informed	**Dependability**	• Percentage of deliveries 'on-time, in-full' • Customers kept informed of delivery dates
• Customisation of terms of insurance cover • Ability to cope with changes in circumstances, such as level of demand • Ability to handle wide variety of risks	**Flexibility**	• Range of possible sizes, gauges, coatings, etc. • Rate of new product introduction • Ability to change quantity, composition and timing of an order
• Premium charged • Arrangement charges • 'No-claims' deals • 'Excess' charges	**Cost**	• Price of products • Price of technical advice • Discounts available • Payment terms

> **OPERATIONS PRINCIPLE**
> ***The interpretation of the five performance objectives will differ between different operations.***

to refer to 'quality of service' as representing all the competitive factors we have listed under quality, speed and dependability (and sometimes aspects of flexibility). For example, information network operations use the term 'quality of service' (QoS) to describe their goal of providing guarantees on the ability of a network to deliver predictable results. This is often specified as including uptime (dependability), bandwidth provision (dependability and flexibility), latency or delay (speed of throughput) and error rate (quality). In practice, the issue is not so much one of universal definition but rather consistency within one, or a group, of operations. At the very least it is important that individual companies have it clear in their own minds how each performance objective is to be defined.

The relative priority of performance objectives differs between businesses

> **OPERATIONS PRINCIPLE**
> ***The relative importance of the five performance objectives depends on how the business competes in its market.***

The idea behind the outside-in perspective is that businesses that compete in different ways should want different things from their operations functions. Therefore, not every operation will apply the same priorities to its performance objectives. There should be a clear logical connection between the competitive stance of a business and its operations objectives. For example, a business that competes primarily on low prices and 'value for money' should place emphasis on operations objectives such as cost, productivity and efficiency; one that competes on a high degree of customisation of its services or products should place an emphasis on flexibility, and so on. Many successful companies understand the importance of making this connection between their message to customers and the operations performance objectives that they emphasise.

Order-winners and qualifiers

> **OPERATIONS PRINCIPLE**
> ***Operations strategy should reflect the requirements of the business's markets.***

A particularly useful way of determining the relative importance of competitive factors is to distinguish between what have been termed 'order-winners' and 'qualifiers'. Figure 2.6 shows the difference between order-winning and qualifying objectives in terms of their utility, or worth, to the competitiveness of the organisation. The curves illustrate the relative amount of competitiveness (or attractiveness to customers) as the operation's performance varies.

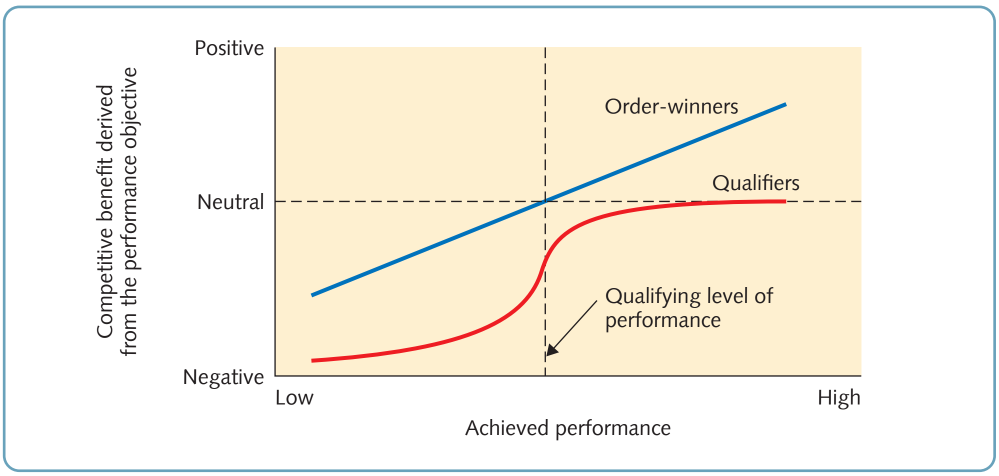


Figure 2.6 Order-winners and qualifiers: order-winners gain more business the better they are performed; qualifiers are the 'givens' of doing business

Order-winners – are those things that directly and significantly contribute to winning business. They are regarded by customers as key reasons for purchasing the product or service. Raising performance in an order-winner will either result in more business or improve the chances of gaining more business. Order-winners show a steady and significant increase in their contribution to competitiveness as the operation gets better at providing them.

Qualifiers – may not be the major competitive determinants of success, but are important in another way. They are those aspects of competitiveness where the operation's performance has to be above a particular level just to be considered by the customer. Performance below this 'qualifying' level of performance may disqualify the operation from being considered by customers. But any further improvement above the qualifying level is unlikely to gain the company much competitive benefit. Qualifiers are those things that are generally expected by customers. Being great at them is unlikely to excite customers, but being bad at them can disadvantage the competitive position of the operation.

Different customer needs imply different objectives

OPERATIONS PRINCIPLE
Different customer needs imply different priorities of performance objectives.

If, as is likely, an operation produces goods or services for more than one customer group, it will need to determine the order-winners and qualifiers for each group. For example, Table 2.3 shows two 'product' groups in the banking industry. Here the distinction is drawn between the customers who are looking for banking services for their private and domestic needs, and the corporate customers who need banking services for their (often large) businesses.

The product/service life cycle influence on performance objectives

One way of generalising the market requirements that operations need to fulfil is to link them to the life cycle of the products or services that the operation is producing. The exact form of

Table 2.3 Different banking services require different performance objectives

Performance objective	*Retail banking*	*Corporate banking*
Products	Personal financial services such as loans and credit cards	Special services for corporate customers
Customers	Individuals	Businesses
Product range	Medium but standardised, little need for special terms	Very wide range, many need to be customised
Design changes	Occasional	Continual
Delivery	Fast decisions	Dependable service
Quality	Means error-free transactions	Means close relationships
Volume per service type	Most services are high volume	Most services are low volume
Profit margins	Most are low to medium, some high	Medium to high
Order-winners	Price Accessibility Speed	Customisation Quality of service Reliability
Qualifiers	Quality Range	Speed Price
Performance objectives emphasised within the processes that produce each service	Cost Speed Quality	Flexibility Quality Dependability

Case example

Dow Corning's operations strategy[3]

(Based on an example from Slack, N. (2017) The Operations Advantage, Kogan Page. Used by permission of Kogan Page.)

For years, Dow Corning was a silicone business with a market position built on service and technical excellence. Customers had been willing to pay top prices for pioneering technology, premium products and customised service. Yet, as the market matured it became clear that some customers were becoming increasingly price sensitive. The premium price strategy was under attack both from large competitors that had driven down costs and from smaller competitors, with lower overheads. Dow Corning was 'stuck in the middle'. In response, Dow Corning decided to undertake detailed analysis of its market and segment customers based on the key factors motivating them to make purchases. Its work revealed four key groups of customers:

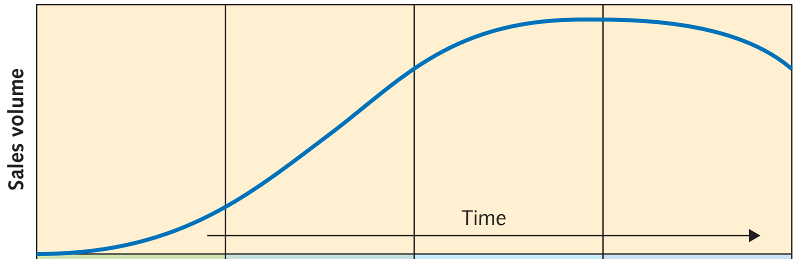
- *Innovative solution seekers* – who wanted innovative silicone-based products.
- *Proven solution seekers* – customers needing advice on existing proven products.
- *Cost-effective solution seekers* – customers who may even pay premium prices for a product, if it could take costs out of their business by improving their productivity.
- *Price seekers* – experienced purchasers of commonly used silicone materials wanting low prices and an easy way of doing business with their supplier.

Each of these segments held a distinct message for Dow Corning's operations. For innovative solution seekers, there was a need to collaborate more closely with customers' R&D personnel in order to develop new products. To target proven solution seekers, the operations function took a more internal approach, working closely with the Dow Corning sales team. The aim was to help them understand its product range in more detail to improve conversion in the sales process. For cost-effective solution seekers, the focus was once again on working closely with the sales staff, but in this case the knowledge transfer was more bi-directional. The key was to build a stronger understanding of customers' processes and help better match their requirements with appropriate offerings. Finally, for the price seeker segment the focus was firmly on bringing down the costs of manufacturing and delivery. This last group was the most challenging for Dow Corning. Its sales to this segment were small and declining, but represented around 30 per cent of the total market for silicones, and was expected to grow significantly. Dow Corning's solution? Create a new offering, called Xiameter. This was a 'no-frills', low-price, restricted range, minimum-order-quantity service, without any technical advice, which could only be accessed on the web (drastically cutting the costs of selling). Delivery times were sufficiently long to fit individual orders into the operation's existing manufacturing schedule.

The development of the Xiameter offering provides a good example of the 'market requirements' perspective on operations strategy. First, Dow Corning segmented the market; this allowed it to identify the differing requirements of the four main customer groups. Second, it decided which segments to serve. Dow Corning decided that while it was weak in the price seeker segment, it was worth pursuing ways in which it might compete. Third, it determined what operations objectives were necessary to compete; for price-seekers, Dow Corning would need to supply at low cost, and abandon its technical advice service. Finally, it determined what operations had to do for Xiameter to be successful: it needed to reduce excess sales overheads (hence web-based sales). Most critically, for this high-volume, low-variety operation to work, customers would need to be prevented from asking for anything that would increase costs (hence limited product range, minimum order quantities and delivery times that would not disrupt production schedules).

product/service life cycles will vary, but generally they are shown as the sales volume passing through four stages – introduction, growth, maturity and decline. The important implication of this for operations management is that products and services will require different operations strategies in each stage of their life cycle (see Figure 2.7).

- *Introduction stage* – when a product or service is first introduced, it is likely to be offering something new in terms of its design or performance. Given the market uncertainty, the operations management of the company needs to develop the flexibility to cope with these changes and the quality to maintain product/service performance.



	Introduction	Growth	Maturity	Decline
	Product/service first introduced to market	Product/service gains market acceptance	Market needs start to be fulfilled	Market needs largely met
Volume	Slow growth in sales	Rapid growth in sales volumes	Sales slow down and level off	Sales decline
Customers	Innovators	Early adopters	Bulk of market	Laggards
Competitors	Few/none	Increasing numbers	Stable number	Declining numbers
Variety of product/service design	Possible high customisation or frequent design changes	Increasingly standardised	Emerging dominant types	Possible move to commodity standardisation
Likely order-winners	Product/service characteristics, performance or novelty	Availability of quality products/services	Low price Dependable supply	Low price
Likely qualifiers	Quality Range	Price Range	Range Quality	Dependable supply
Dominant process performance objectives	Flexibility Quality	Speed Dependability Quality	Cost Dependability	Cost

Figure 2.7 The effects of the product/service life cycle on the operation and its process performance objectives

- *Growth stage* – in the growing market, standardised designs emerge that allow the operation to supply the rapidly growing market. Keeping up with demand through rapid and dependable response and maintaining quality levels will help to keep market share as competition starts to increase.
- *Maturity stage* – eventually demand starts to level off as the market becomes dominated by a few larger companies with standardised designs. Competition will probably emphasise price or value for money, so operations will be expected to get the costs down in order to maintain profits or to allow price-cutting, or both. So, cost and productivity issues, together with dependable supply, are likely to be the operation's main concerns.
- *Decline stage* – after time, sales will decline. To the companies left there might be a residual market, but if capacity in the industry lags demand, the market will be dominated by price competition; therefore cost-cutting continues to be important.

Worked example

Innovation at Micraytech (Part 2, outside-in)

The Micray Group sees a major growth opportunity for Micraytech by continually incorporating technological innovations in its product offerings. However, Micraytech's marketing management know that this can be achieved by focusing on one or both of two distinct markets. The first is the market for 'individual metrology devices'. These are 'stand-alone' pieces of equipment bought by all types of industrial customers. It had traditionally been the company's main market. The second market is that for 'integrated metrology systems'. These are larger, more complex, more expensive (and higher-margin) offerings that are customised to individual customers' requirements. The two types of offering have overlapping but different characteristics. 'Individual metrology devices' compete on their technical performance and reliability, together with relatively short delivery times compared with competitors. The 'integrated metrology systems' offerings currently account for only a small part of the company's sales, but it is a market that is forecast to grow substantially. The customers for these systems are larger manufacturers who are investing in more automated technologies and require metrology systems that can be integrated into their processes. From an 'outside-in' perspective, if it was to take advantage of this emerging market, Micraytech would have to learn how to work more closely with both its direct customers and the firms that were supplying their customers with the automated technologies. In addition to Micraytech's traditional technical skills, it would have to increase its software development, data exchange and client liaison skills.

2.4 Diagnostic question: *Does operations strategy learn from operational experience (bottom-up)?*

Although it is a convenient way of thinking about strategy, the top-down hierarchical model does not always represent the way strategies are formulated in practice. When any group is reviewing its corporate strategy, it will also take into account the circumstances, experiences and capabilities of the various businesses that form the group. This is sensible. Operations strategy should always reflect operational reality. So, businesses, when reviewing their strategies, should consult the individual functions within the business about their constraints and capabilities. They may also incorporate the ideas that come from each function's day-to-day experience. In fact, many strategic ideas emerge over time from operational experience rather than originating exclusively at a senior level.

Sometimes companies move in a particular strategic direction because the ongoing experience of providing products and services to customers at an operational level convinces them that it is the right thing to do. There may be no formal high-level decision-making that examines alternative strategic options and chooses the one that provides the best way forward. Instead, a general consensus emerges from the operational experience. The 'high-level' strategic decision-making, if it occurs at all, may simply confirm the consensus and provide the resources to make it happen effectively. This is sometimes called the concept of 'emergent strategies'. It sees strategy-making, at least partly, as a relatively unstructured and fragmented process to reflect the fact that the future is at least partially unknown and unpredictable.

> **OPERATIONS PRINCIPLE**
> ***Operations strategy should reflect the bottom-up experience of operational reality.***

This perspective on operations strategy reflects how things often happen, but at first glance it seems less useful in providing a guide for specific decision-making. Yet, while emergent strategies are less easy to categorise, the principle governing a bottom-up perspective is clear: operations' objectives and actions should be shaped, at least partly, by the knowledge it gains from its day-to-day activities. Certainly, encouraging strategic ideas to emerge from the operational experience is not an abandonment of strategic responsibility by senior management. Rather it

Worked example

Innovation at Micraytech (Part 3, bottom-up)

Over time, as its operations strategy developed, Micraytech discovered that continual product and system innovation was having the effect of dramatically increasing its costs. And, although it was not competing on low prices, and nor was it under pressure from the Group to achieve high rates of return on sales, its rising costs were impacting profitability to an unacceptable degree. There was also some evidence that continual updating of product and system specifications was confusing some customers. Partially in response to customer requests, the company's system designers started to work out a way of 'modularising' their system and product designs. This allowed one part of the system to be updated for those customers who valued the functionality that the innovation could bring, without interfering with the overall design of the main body of the system, of which the module was a part. Over time, this approach became standard design practice within the company. Customers appreciated the extra customisation, and modularisation reduced operations costs. Note that this strategy emerged from the company's experience. It was a pure bottom-up approach. Initially, no top-level board decision was taken to initiate this practice. Nevertheless, it emerged as the way in which the company's design engineers learned from their experience and used that learning to build their knowledge of how to lower some of the costs of innovation.

is an acceptance that worthwhile ideas can come from those who work at the operational level of the business. In fact, not accepting it would be to waste one of the most insightful sources of practical understanding of operational reality. From a bottom-up perspective, the key virtue required for shaping strategy is an ability to learn from experience. More specifically, the bottom-up perspective should involve:

- capturing the learning that comes from routine operations activities; and
- transforming that learning into strategically valuable knowledge.

Another advantage of paying attention to the bottom-up perspective is that it is at the operational day-to-day level of experience where trends often first become evident, and for businesses operating in unstable or unpredictable environments this can be particularly important. The other three perspectives of operations strategy can take time to detect trends in how markets are moving. The bottom-up element is more 'plugged-in' to everyday experience.

2.5 Diagnostic question: *Does operations strategy develop the capability of its resources and processes (inside-out)?*

Operations strategy must do more than simply meet the short-term needs of the market (important though this is). The processes and resources within operations also need to be developed in the long term to provide the business with a set of competences or capabilities (we use the two words interchangeably). A capability in this context is the 'know-how' that is embedded within the business's resources and processes. These capabilities may be built up over time as the result of the experiences of the operation (bottom-up), or they may be bought-in or acquired. If they are refined and integrated, they can form the basis of the business's ability to offer unique and/or 'difficult-to-imitate' products and services to its customers. This idea of the basis of long-term competitive capabilities deriving from the operation's resources and processes is called the resource, or inside-out, perspective on operations strategy.

Operations strategy should build operations capabilities

Building operations capabilities means understanding the existing resources and processes within the operation, starting with the simple questions what do we have and what can we do? However, trying to understand an operation by listing its resources alone is like trying to understand an automobile by listing its component parts. To understand an automobile we need to describe how the component parts form its internal mechanisms. Within the operation, the equivalents of these mechanisms are its processes. Yet, even a technical explanation of an automobile's mechanisms does not convey its style or 'personality'. Something more is needed to describe these. In the same way, an operation is not just the sum of its processes. It also has intangible resources. An operation's intangible resources include such things as:

- its relationship with suppliers and the reputation it has with its customers;
- its knowledge of and experience in handling its process technologies;
- the way its staff can work together in new product and service development;
- the way it integrates all its processes into a mutually supporting whole.

> **OPERATIONS PRINCIPLE**
> ***The long-term objective of operations strategy is to build operations-based capabilities.***

These intangible resources may not be as evident within an operation, but they are important and often have real value. And both tangible and intangible resources and processes shape its capabilities. The central issue for operations management, therefore, is to ensure that its pattern of strategic decisions really does develop appropriate capabilities.

An operations strategy should identify the broad decisions that will help the operation build its capabilities

Few businesses have the resources to pursue every single action that might improve their operations performance. So, an operations strategy should indicate broadly how the operation might best achieve its performance objectives. For example, a business might specify that it will attempt to reduce its costs by aggressive outsourcing of its non-core business processes and by investing in more efficient technology. Or, it may declare that it intends to offer a more customised set of products or services through adopting a modular approach to its product or service design. The balance here is between a strategy that is overly restrictive in specifying how performance objectives are to be achieved, and one that is so open that it gives little guidance as to what ideas should be pursued.

There are several categorisations of operations strategy decisions. Any of them are valid if they capture the key decisions. Table 2.4 illustrates some of the broad operations strategy decisions that fall within each category.

The resource-based view

The idea that building operations capabilities is a particularly important objective of operations strategy is closely linked with the 'resource-based view' (RBV) of the firm. This holds that businesses with an 'above-average' strategic performance are likely to have gained their sustainable competitive advantage because of their core competences (or capabilities). This means that the way an organisation inherits, or acquires, or develops its operations resources will, over the long term, have a significant impact on its strategic success. The RBV differs in its approach from the more traditional view of strategy that sees companies as seeking to protect their competitive advantage through their control of the market. For example, they may do this by creating barriers to entry through product differentiation, or making it difficult for customers to switch to competitors, or controlling the access to distribution channels (a major barrier to

Table 2.4 Some strategic decisions that may be addressed in an operations strategy

- How should the operation decide which products or services to develop and how to manage the development process?
- Should the operation outsource some of its activities, or take more activities in-house?
- Should the operation expand by acquiring its suppliers or its customers? If so, which ones should it acquire?
- How many geographically separate sites should the operation have?
- Where should operations sites be located?
- What activities and capacity should be allocated to each site?
- What broad types of technology should the operation be using?
- How should the operation be developing its people?
- What role should the people who staff the operation play in its management?
- How should the operation forecast and monitor the demand for its offerings?
- How should the operation adjust its activity levels in response to demand fluctuations?
- How should the operation monitor and develop its relationship with its suppliers?
- How much inventory should the operation have and where should it be located?
- What system should the operation use to coordinate its activities?
- How should the operation's performance be measured and reported?
- How should the operation ensure that its performance is reflected in its improvement priorities?
- Who should be involved in the improvement process?
- Should improvement be continuous, or radical, or both?
- How should the improvement process be managed?
- How should the operation maintain its resources so as to prevent failure?
- How should the operation ensure continuity if a failure occurs?

entry in gasoline retailing, for example, where oil companies own their own retail stations). By contrast, the RBV sees firms being able to protect their competitive advantage through barriers to imitation – that is, by building up 'difficult-to-imitate' resources. Some of these 'difficult-to-imitate' resources are particularly important, and can be classified as 'strategic' if they exhibit the following properties:

- They are scarce – scarce resources, such as specialised production facilities, experienced engineers or proprietary software, can underpin competitive advantage.
- They are imperfectly mobile – some resources are difficult to move out of a firm. For example, resources that were developed in-house, or are based on the experience of the company's staff, or are interconnected with the other resources in the firm, cannot be traded easily.
- They are imperfectly imitable and imperfectly substitutable – it is not enough only to have resources that are unique and immobile. If a competitor can copy these resources, or replace them with alternative resources, then their value will quickly deteriorate. The more the resources are connected with process knowledge embedded deep within the firm, the more difficult they are for competitors to understand and to copy.

The VRIO framework

The most common (and useful) way of applying the RBV has become known as the VRIO framework. It was first developed by Jay Barney in the 1990s (who originally identified the idea of resources needing to be scarce, imperfectly mobile, imperfectly imitable and imperfectly substitutable) but later modified to make it more useful for practitioners. In this framework, the four questions to ask about any potentially strategic resource are:

1. ***Is the resource valuable?*** Is it possible to identify specific and definable competitive value from the resource?

2. ***Is the resource rare?*** Do you have (or have access to) resources that your competitors do not? Some theorists define the idea of 'rarity' as when a business has a resource that is unequivocally unique, but for all practical purposes a resource is 'rare' if it is, at least, in short supply and likely to remain so.

3. ***Is the resource costly to imitate?*** Do you have resources that competitors cannot imitate, purchase or find a suitable alternative to, at a realistic cost or in a realistic time frame? 'Imitability' can be either because competitors can copy your resources and processes directly, or because they can find an acceptable substitute for them.

4. ***Is the firm organised to capture the value of the resource?*** Do you have within your business the systems, culture, capacity and motivation to exploit any capabilities embedded in your resources and processes? A firm must have the formal reporting and control mechanisms, leadership, and the informal and cultural environment that allows the strategic resources to develop.

There are two important points to remember about the VRIO framework. First, all these factors are time dependent. A capability may be valuable now, but competitors are unlikely to stand still. In addition, rarity and inimitability are not absolutes and, with time, can be undermined by competitor activity. Even the ability to exploit capabilities can erode if operations leadership is lacking. Second, although the conventional order in which to treat each of these elements is as we have done here (which is why it's called the VRIO framework), it maybe is best to think of the 'O' of 'organisation' as a necessary prerequisite. Without the ability to exploit strategic resources, they are of little use. However, with effective organisation there is the potential for operations resources to contribute to competitiveness. If their capabilities are also valuable, then parity with competitors should be possible. With the addition of rarity, a short- to medium-term competitive advantage is possible. With the addition of inimitability, competitors will find it difficult to match capabilities in anything but the long term. This sequence is shown in Figure 2.8.

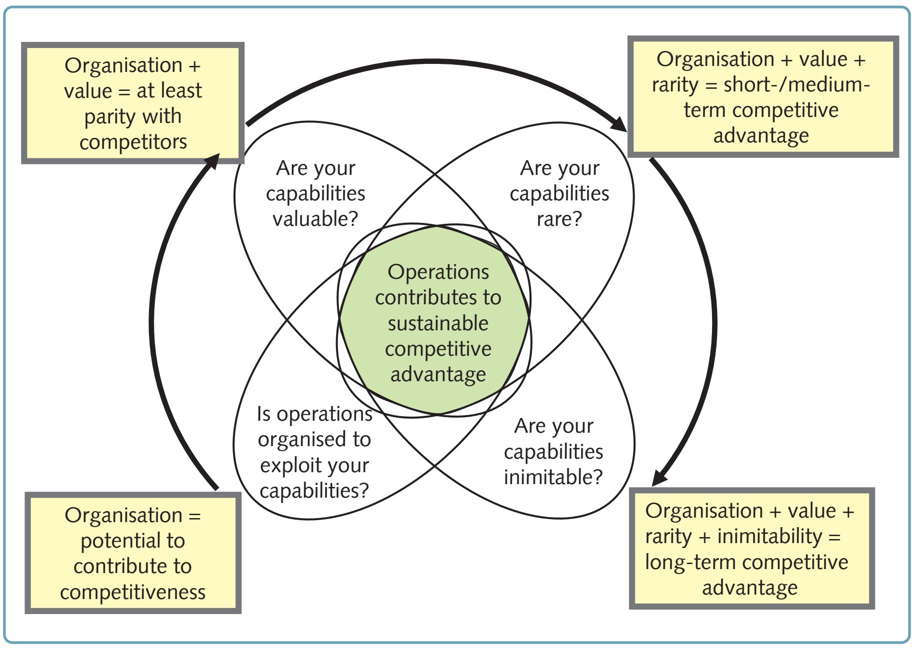


Figure 2.8 The four features of Barney's VRIO framework

Source: From Slack, N. (2017) *The Operations Advantage*, Kogan Page, reproduced by permission.

Worked example

Innovation at Micraytech (Part 4, inside-out)

The modular approach to product design proved to be a big success for Micraytech. However, it posed two challenges for the company's operations. First, the technical aspects of integrating some of the more sophisticated modules proved difficult. This affected only a small proportion of customers, but they were the ones that were willing to pay premium prices for their systems. The only potential solution was to attempt to develop the interface modules that would allow previously incompatible modules to be integrated. When this solution was first proposed the relevant skills were not present in the company. It had to recruit specialist engineers to start on the design of the interfaces. During this design process the company realised that it could potentially open up a new market. As the firm's chief operating officer (COO) put it, *'If we designed the interfaces carefully, we could not only integrate all of our own in-house modules, we could also integrate other firms' instruments into our systems'.* This led to the second set of challenges: to develop relationships with possible suppliers, who might very well be competitors in some markets, so that they were willing to supply their equipment for inclusion in Micraytech's systems. Not only this, but the firm also had to ensure that the internal processes, of its sales engineers consulting with clients, its design department designing the system to clients' needs, and its procurement managers negotiating with equipment suppliers, all operated seamlessly. *'The success that we have enjoyed can be put down to two key capabilities. The first was to buy-in the engineering skills to create technically difficult interfaces. That led to us understanding the value that could be gained from a seamless internal and external supply chain. Both of these capabilities are not totally impossible for other firms to copy, but they would be very difficult for them to get to our level of excellence.'* (COO Micraytech)

2.6 Diagnostic question: *Are the four perspectives of operations strategy reconciled?*

As we stressed earlier, none of the four perspectives alone can give a full picture of any organisation's operations strategy. But together they do provide a good idea of how its operations are contributing strategically. For example, Figure 2.9 brings together the four perspectives of the Micraytech operations strategy. For Micraytech, the four perspectives seem to be reasonably compatible, with its operations strategy fitting together from whichever perspective is chosen. In other words, each perspective is 'reconciled' with the others. This is one of the conditions for an effective operations strategy – the four perspectives must be reconciled.

The operations strategy matrix

OPERATIONS PRINCIPLE
An operations strategy should articulate the relationship between operations objectives and the means of achieving them.

The operations strategy matrix is one method of checking the reconciliation between the inside-out and outside-in perspectives. It brings together (a) market requirements and (b) operations resources to form the two dimensions of a matrix. It describes operations strategy as the intersection of a company's performance objectives and the strategic decisions that it makes. In fact, there are several intersections between each performance objective and each decision area (however one wishes to define them). If a business thinks that it has an operations strategy, then it should have a coherent explanation for each of the cells in the matrix. That is, it should be able to explain and reconcile the intended links between each performance objective and each decision area. The process of reconciliation takes place between what is required from the operations function (performance objectives), and how the operation tries to achieve this through the set of choices made (and the capabilities that have been developed) in each decision area.

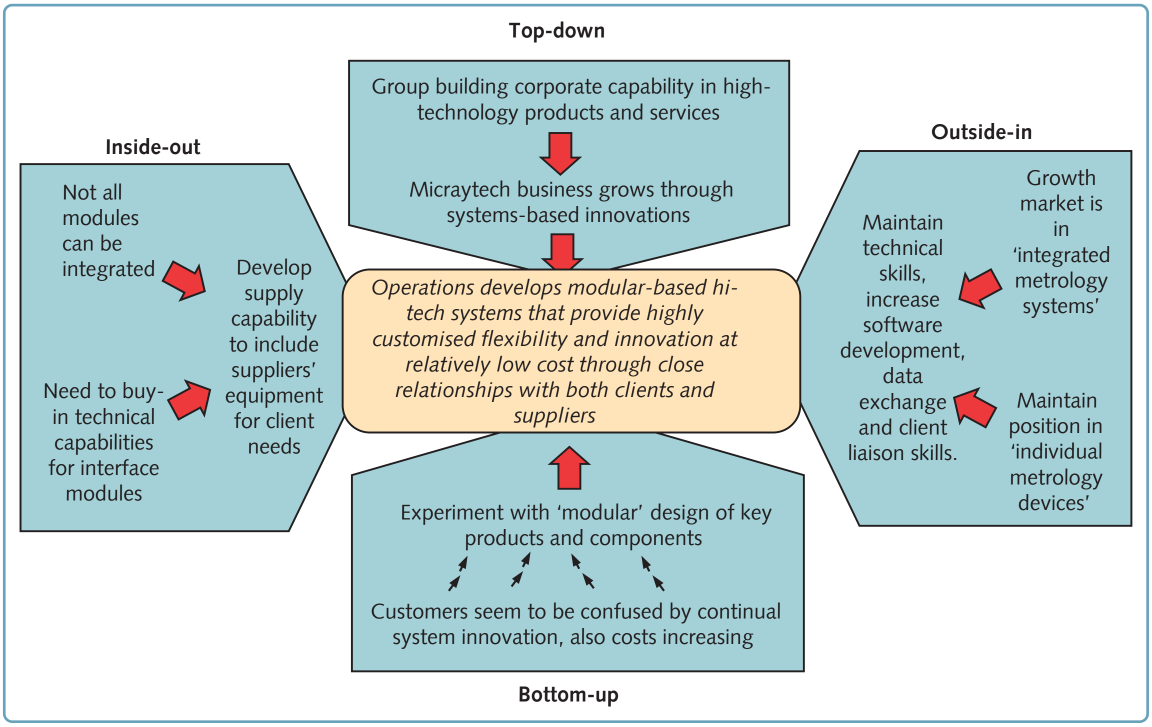


Figure 2.9 **Top-down, outside-in, bottom-up and inside-out perspectives of the Micraytech operations strategy**

Figure 2.10 shows a simplified example of how the operations strategy matrix can be used. A parcel courier service competes primarily on its quality and dependability of service, with price (cost) and innovation also being fairly important. The range of services offered is not unimportant, but not of prime concern. Figure 2.10 illustrates that the company believes its quality of service is going to be influenced largely by investment in 'track and trace' technology (which allows customers to check where their delivery is) and its knowledge management system (which allows improvements in its processes to be recorded and shared). Other key intersections are as illustrated in the figure. Note that not all cells are occupied. This is not because there is no relationship between the performance objectives and the decisions that these cells represent, it is rather that the decisions are not seen as being particularly important in the context of the whole strategy.

Reconciling market requirements and operations capabilities over time – the 'line of fit' model

The operations strategy matrix is a good model for testing whether market requirements and the operations capability perspectives fit together. It makes explicit the specific aspects of market requirements (quality, speed, dependability, flexibility, cost, etc.) and the decisions that support operations capability (design, delivery and development). The disadvantage is that it gives little sense of the dynamics of reconciliation – how the balance between market requirements and the operations capability changes over time. This is where the 'line of fit' model is useful. It is based on the idea that, ideally, there should be a reasonable degree of alignment, or 'fit', between the requirements of the market and the capabilities of the operation. Figure 2.11 illustrates the concept of fit diagrammatically. The vertical dimension represents the (outside-in) nature of market requirements either because they reflect the intrinsic needs of customers, or because their expectations have been shaped by the firm's marketing activity. This includes such factors as the strength of the brand, or reputation of the degree of differentiation, or the

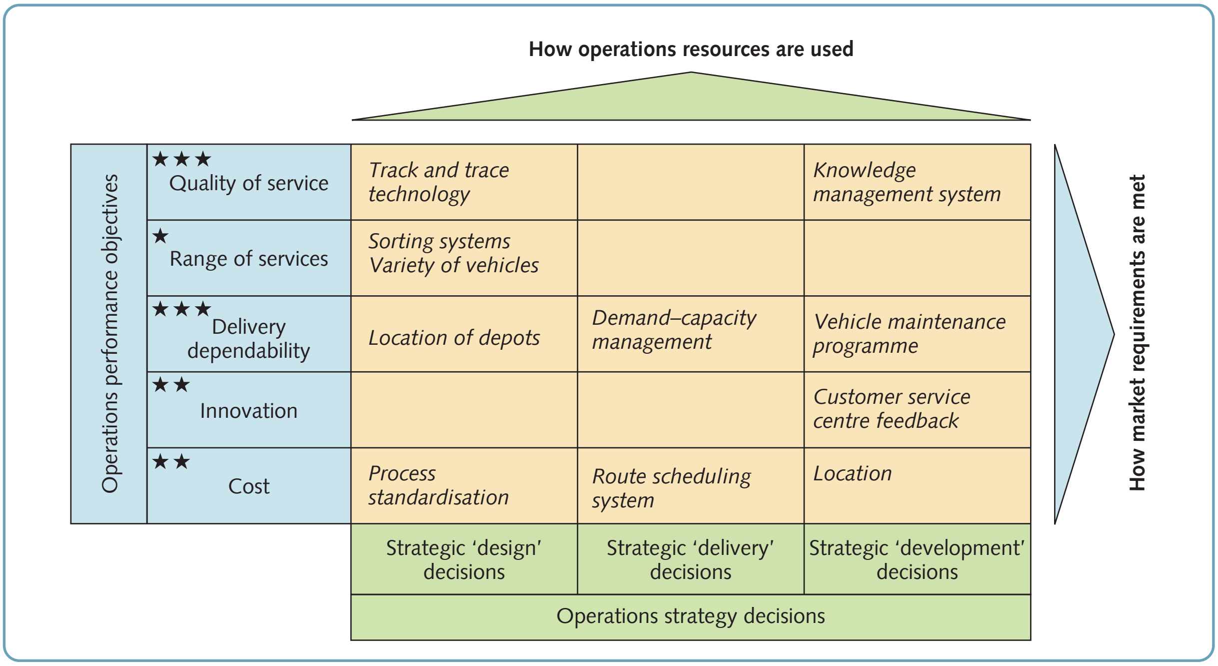


Figure 2.10 The operations strategy matrix defines operations strategy by the intersections of performance objectives and operations decisions – in this case for a parcel delivery courier

extent of market promises. Movement along the dimension indicates a broadly enhanced level of market 'performance'. The horizontal scale represents the (inside-out) nature of the firm's operations resource and process capability. This includes such things as the performance of the operation in terms of its ability to achieve competitive objectives, the efficiency with which it uses its resources and the ability of the firm's resources to underpin its business. Movement along the dimension broadly indicates an enhanced level of 'operations capability'.

If the market requirements and operations capability of an operation are aligned they would be positioned diagrammatically on the 'line of fit' in Figure 2.11. 'Fit' is to achieve an approximate balance between 'market requirements' and 'operations capability'. So when fit is achieved, firms' customers do not need, or expect, levels of operations capability that cannot be supplied. Nor does the operation have strengths that are either inappropriate for market needs or remain unexploited in the market.

> **OPERATIONS PRINCIPLE**
> ***An operation positioned on the 'line of fit' has operations capabilities that match its market requirements.***

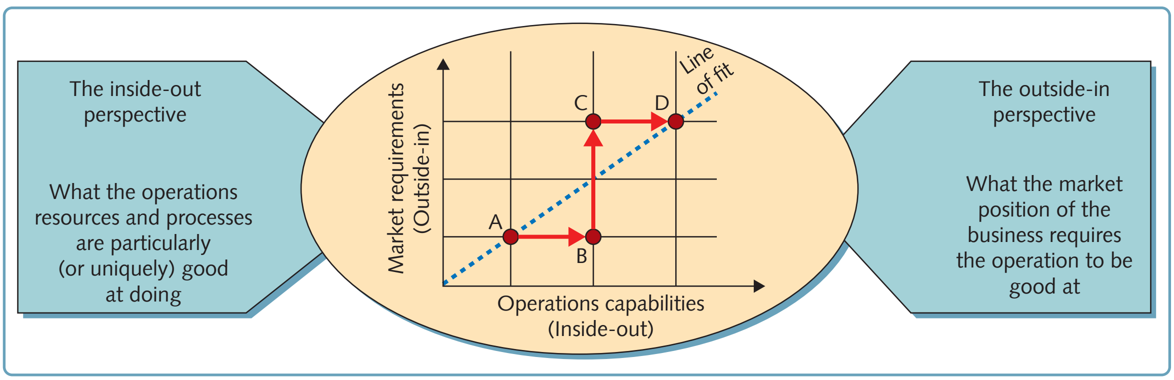


Figure 2.11 The 'line of fit' model shows how operations strategy attempts to reconcile market requirements and operations capabilities over time

A company that has position A in Figure 2.11 has achieved 'fit' in so much as its operations capabilities are aligned with its market requirements, yet both are at a relatively low level. In other words, the market does not want much from the business, which is just as well because its operation is not capable of achieving much. Over time its ambition is to move to position D, where it has also achieved 'fit', but at a much higher level. Other things being equal, this will be a more profitable position than position A.

However, like most strategic improvement, the company cannot always guarantee to keep on the 'line of fit' as it moves from A to D over time. In this case, it first improves its operations capability without exploiting its enhanced capability in its market (position B). This could be seen as a 'waste' of its potential to adopt a more ambitious (and possibly profitable) market position. Realising this, the company revises its marketing strategy to promote itself as being able to maintain a much higher level of market performance (position C). Unfortunately, these new promises to its market are not matched by its operations capabilities. The company is again away from the line of fit. In fact, position C is possibly even more damaging than position B. The risk now is that the company's market reputation will erode until it can improve its operations capabilities to bring it back to the line of fit (position D).

The issue that is highlighted by positioning operations strategy relative to the line of fit is that progress cannot always be a smooth trajectory that achieves perfect balance between market requirements and operations capability. Furthermore, when an operation deviates from the line of fit there are predictable consequences. A position below the line of fit means that the operation is failing to exploit its operations capabilities. A position above the line of fit means that it risks damaging its reputation or brand by failing to live up to its market promises.

2.7 Diagnostic question: *Does operations strategy set an improvement path?*

An operations strategy is the starting point for operations improvement. It sets the direction in which the operation will change over time. It is implicit that the business will want operations to change for the better. Therefore, unless an operations strategy gives some idea as to how improvement will happen, it is not fulfilling its main purpose. This is best thought about in terms of how performance objectives, both in themselves and relative to each other, will change over time. To do this, we need to understand the concept of, and the arguments concerning, the trade-offs between performance objectives.

An operations strategy should guide the trade-offs between performance objectives

An operations strategy should address the relative priority of operations performance objectives ('for us, speed of response is more important than cost efficiency, quality is more important than variety', and so on). To do this it must consider the possibility of improving its performance in one objective by sacrificing performance in another. So, for example, an operation might wish to improve its cost efficiencies by reducing the variety of products or services that it offers to its customers. Taken to its extreme; this 'trade-off' principle implies that improvement in one performance objective can only be gained at the expense of another. 'There is no such thing as a free lunch' could be taken as a summary of this approach to managing. Probably the best-known summary of the trade-off idea comes from Professor Wickham Skinner, the most influential of the originators of the strategic approach to operations, who said:[4]

'. . . most managers will readily admit that there are compromises or trade-offs to be made in designing an airplane or truck. In the case of an airplane, trade-offs would involve matters such as cruising speed, take-off and landing distances, initial cost, maintenance, fuel consumption, passenger comfort and cargo or passenger capacity. For instance, no one today can design a 500-passenger plane that can land on an aircraft carrier and also break the sound barrier. Much the same thing is true in . . . [operations].'

OPERATIONS PRINCIPLE
In the short term, operations cannot achieve outstanding performance in all its operations objectives simultaneously.

OPERATIONS PRINCIPLE
In the long term, a key objective of operations strategy is to improve all aspects of operations performance.

But there is another view of the trade-offs between performance objectives. This sees the very idea of trade-offs as the enemy of operations improvement, and regards the acceptance that one type of performance can only be achieved at the expense of another as both limiting and unambitious. For any real improvement of total performance, it holds, the effect of trade-offs must be overcome in some way. In fact, overcoming trade-offs must be seen as the central objective of strategic operations improvement.

These two approaches to managing trade-offs result in two approaches to operations improvement. The first approach emphasises 'repositioning' performance objectives by trading-off improvements in some objectives for a reduction in performance in others; the second one emphasises increasing the 'effectiveness' of the operation by overcoming trade-offs so that improvements in one or more aspects of performance can be achieved without any reduction in the performance of others. Most businesses at some time or other will adopt both approaches. This is best illustrated through the concept of the 'efficient frontier' of operations performance.

Case example

Tesco turns itself around[5]

When market conditions change, it usually means that operations strategy should change. But it can take time before changes become obvious, and even longer to react. This was a lesson that Tesco, the UK's biggest food retailer, learned when, after years of profitable growth, and although still the market leader in the UK, its lead over rivals worsened significantly. Why, asked its detractors, had the company not realised that its strategy was failing? Some analysts pointed out that it was partly because of competitor activity. Waitrose (an upmarket supermarket, with a good reputation for quality) was serving the top end of the market, while German discount stores Aldi and Lidl were attracting more cost-conscious customers. However, it was also pointed out that some problems were of Tesco's own making, caused by its operations strategy failing to respond fast enough to market requirements. The strategy of building large out-of-town superstores had continued, even though a sharper monitoring of consumer behaviour would have revealed that such large-capacity units had lost their attraction as families cut down on weekly trips to the supermarket and opted instead for home deliveries while topping up their groceries with trips to local stores. In fact, Philip Clarke, Tesco's then Chief Executive, admitted that he ought to have moved faster to cut back on planned superstore openings in response to clear radical changes in shopping habits.

This episode in Tesco's (largely successful) history provided important lessons. Making significant change in operations strategy can be extremely disruptive and costly in the short term. A new Chief Executive, Dave Lewis, halted the store expansion programme and 43 branches (out of a UK total of about 4,000) were closed. The savings from this were reinvested into improved

jax10289/Shutterstock

▶

customer service. Tesco's relationships with its suppliers, which had been described as toxic, were overhauled. largely by empowering existing teams to make better decisions. The directive was that the suppliers and Tesco should imagine they were all one company and work out together the most efficient supply chain. The turnaround worked. When the company's trading stabilised, Tesco was able to focus on other priorities. One of these was sustainability: Tesco became a pioneer in reporting and measuring food waste.

Trade-offs and the 'efficient frontier'

Figure 2.12(a) shows the relative performance of several companies in the same industry in terms of their cost efficiency and the variety of products or services that they offer to their customers. Presumably all the operations would ideally like to be able to offer very high variety, while still having very high levels of cost efficiency. However, the increased complexity that a high variety of product or service offerings brings will generally reduce the operation's ability to operate efficiently. Conversely, one way of improving cost efficiency is to severely limit the variety on offer to customers. The spread of results in Figure 2.12(a) is typical of an exercise such as this. Operations A, B, C and D all have chosen a different balance between variety and cost efficiency. But none is dominated by any other operation in the sense that another operation necessarily has 'superior' performance. Operation X, however, has an inferior performance because operation A is able to offer higher variety at the same level of cost efficiency; and operation C offers the same variety but with better cost efficiency. The convex line on which operations A, B, C and D lie is known as the 'efficient frontier'. They may choose to position themselves differently (presumably because of different market strategies) but they cannot be criticised for being ineffective. Of course, any of these operations that lie on the efficient frontier may come to believe that the balance they have chosen between variety and cost efficiency is inappropriate. In these circumstances they may choose to reposition themselves at some other point along the efficient frontier. By contrast, operation X has also chosen to balance variety and cost efficiency in a particular way but is not doing so effectively. Operation B

OPERATIONS PRINCIPLE
Operations that lie on the 'efficient frontier' have performance levels that dominate those that do not.

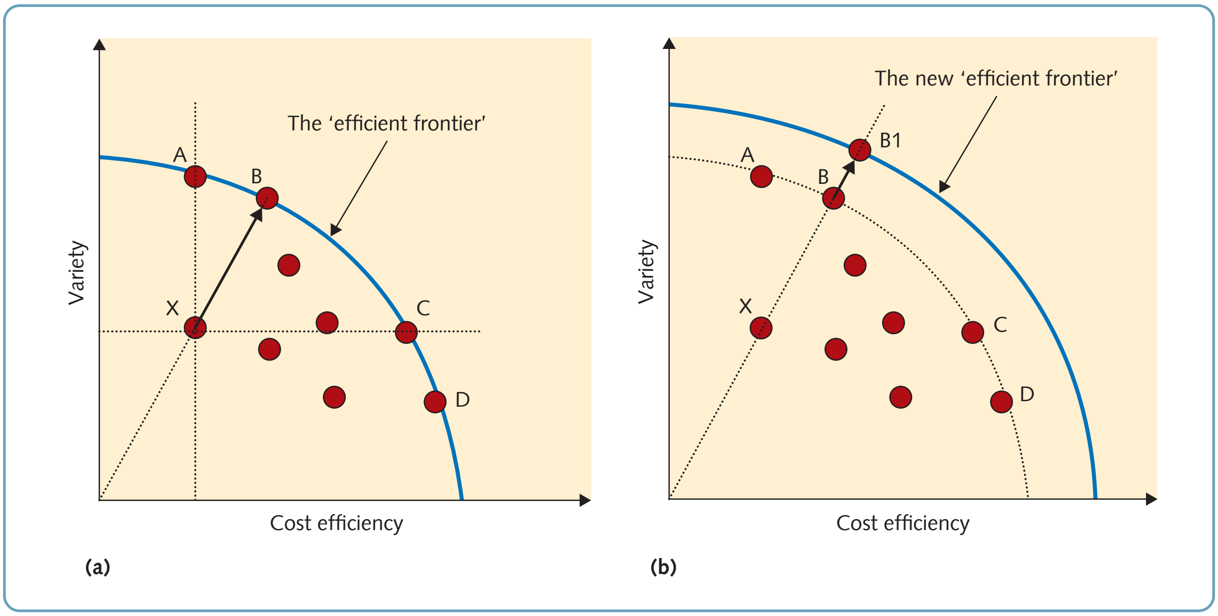


Figure 2.12 The 'efficient frontier'

has the same ratio between the two performance objectives but is achieving them more effectively. Operation X will generally have a strategy that emphasises increasing its effectiveness before considering any repositioning.

However, a strategy that emphasises increasing effectiveness is not confined to those operations that are dominated, such as operation X. Those with a position on the efficient frontier will generally also want to improve their operation's effectiveness by overcoming the trade-off that is implicit in the efficient frontier curve. For example, suppose operation B in Figure 2.12(b) is the metrology company described earlier in this chapter. By adopting a modular product design strategy, it improved both its variety and its cost efficiency simultaneously (and moved to position B1). What has happened is that operation B has adopted a particular operations practice (modular design) that has pushed out the efficient frontier. This distinction between positioning on the efficient frontier and increasing operations effectiveness to reach the frontier is an important one. Any operations strategy must make clear the extent to which it is expecting the operation to reposition itself in terms of its performance objectives and the extent to which it is expecting the operation to improve its effectiveness.

Improving operations effectiveness by using trade-offs

Improving the effectiveness of an operation by pushing out the efficient frontier requires different approaches, depending on the original position of the operation on the frontier. For example, in Figure 2.13(a) operation P is originally resourced and designed to offer a high level of variety of products or services at the expense of low-cost efficiency. It has probably reached this position by adopting a series of operations practices that enable it to offer the variety, even if these practices are intrinsically expensive. For example, it may have invested in general-purpose technology and recruited employees with a wide range of skills. Improving variety even further may mean adopting more extreme operations practices that emphasise variety. For instance, it may reorganise its processes so that each of its larger customers has a dedicated set of resources that understands the specific requirements of that customer and can organise itself to totally customise every product or service it produces. This will probably mean a further sacrifice of cost efficiency, but it allows an ever-greater variety of products or services to

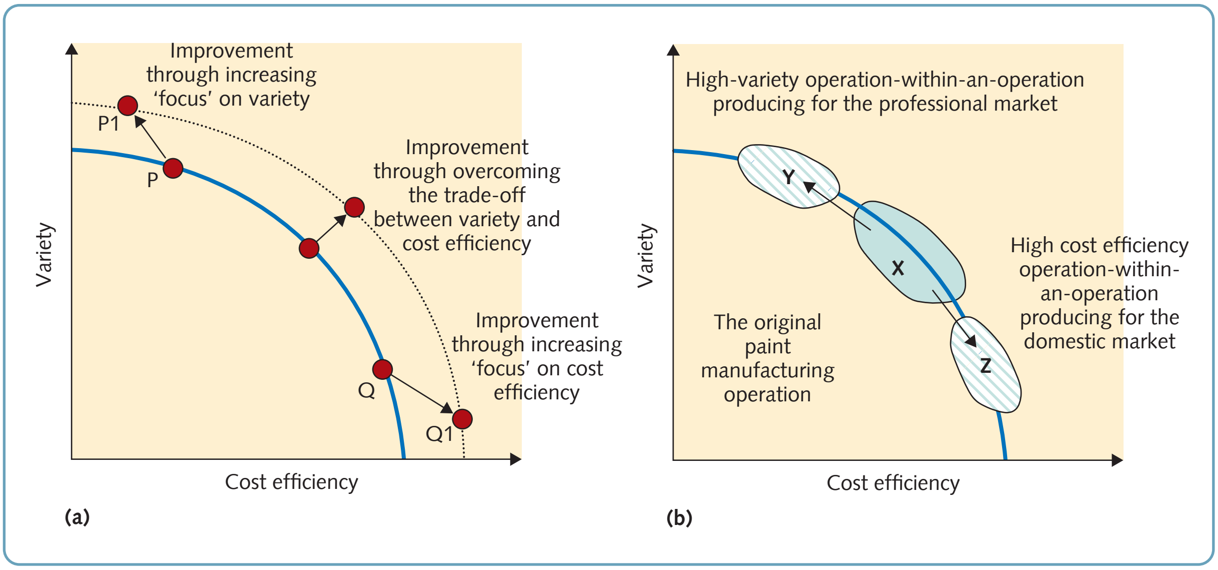


Figure 2.13 Operations 'focus' and the 'operation-within-an-operation' concept illustrated using the 'efficient frontier' model

be produced (P1). Similarly, operation Q may increase the effectiveness of its cost efficiency, by becoming even less able to offer any kind of variety (Q1). For both operations P and Q, effectiveness is being improved through increasing the focus of the operation on a very narrow set of performance objectives and accepting an even further reduction in other aspects of performance.

The same principle of focus also applies to organisational units smaller than a whole operation. For example, individual processes may choose to position themselves on a highly focused set of performance objectives that match the market requirements of their own customers. So, for example, a business that manufactures paint for interior decoration may serve two quite distinct markets. Some of its products are intended for domestic customers who are price-sensitive but demand only a limited variety of colours and sizes. The other market is professional interior decorators who demand a very wide variety of colours and sizes but are less price-sensitive. The business may choose to move from a position where all types of paint are made on the same processes (position X in Figure 2.13(b)) to one where it has two separate sets of processes (positions Y and Z) – one that only makes paint for the domestic market and the other that only makes paint for the professional market. In effect, the business has segmented its operations processes to match the segmentation of the market. This is called the 'operation-within-an-operation' (or 'plant-within-a-plant', 'shop-within-a-shop', etc.) concept.

Improving operations effectiveness by overcoming trade-offs

This concept of highly focused operations is not universally seen as appropriate. Many companies attempt to give 'the best of both worlds' to their customers. At one time, for example, a high-quality, reliable and error-free automobile was inevitably an expensive automobile. Now, with few exceptions, we expect even budget-priced automobiles to be reliable and almost free of any defects. Auto manufacturers found that not only could they reduce the number of defects on their vehicles without necessarily incurring extra costs, but also they could actually reduce costs by reducing errors in manufacture. If auto manufacturers had adopted a purely focus-based approach to improvement over the years, we may now only be able to purchase either very cheap low-quality automobiles or very expensive high-quality automobiles. So, a permanent expansion of the efficient frontier is best achieved by overcoming trade-offs through improvements in operations practice.

> OPERATIONS PRINCIPLE
> *An operation's strategy improvement path can be described in terms of repositioning and/or overcoming its performance trade-offs.*

Even trade-offs that seem to be inevitable can be reduced to some extent. For example, one of the decisions that any supermarket manager has to make is how many checkout positions to open at any time. If too many checkouts are opened then there will be times when the checkout staff do not have any customers to serve and will be idle. The customers, however, will have excellent service in terms of little or no waiting time. Conversely, if too few checkouts are opened, the staff will be working all the time but customers will have to wait in long queues. There seems to be a direct trade-off between staff utilisation (and therefore cost) and customer waiting time (speed of service). Yet even the supermarket manager deciding how many checkouts to open can go some way to affecting the trade-off between customer waiting time and staff utilisation. The manager might, for example, allocate a number of 'core' staff to operate the checkouts but also arrange for those other staff who are performing other jobs in the supermarket to be trained and 'on-call' should demand suddenly increase. If the manager on duty sees a build-up of customers at the checkouts, these other staff could quickly be used to staff checkouts. By devising a flexible system of staff allocation, the manager can both improve customer service and keep staff utilisation high.

Critical commentary

- Starting any discussion of strategy from a stakeholder perspective is far from undisputed. Using so many criteria is too complex. If one needs to dictate the criterion to be maximised by operations, what exactly would it be? In other words, how do we want performance to be measured? How do we want operations to determine what is worthwhile versus what is not? It could also be argued that using stakeholder perspectives gives undue weight to narrow special interests who want to use the organisation's resources for their own ends. The stakeholder perspective gives them a spurious legitimacy, which undermines the core importance of value-seeking behaviour.
- Similarly, the idea that operations strategy could ever become the driver of a business's overall strategy, and the associated concept of the resource-based view of the firm, are both problematic to some theorists. Business strategies and functional strategies were, for many years, seen as first market driven and second planned in a systematic and deliberative manner. So, it became almost axiomatic to see strategy as starting from a full understanding of market positioning. In fact, the main source of sustainable competitive advantage was seen as unequivocally associated with how a business positioned itself in its markets. Get the market proposition right and customers would respond by giving you business. Get it wrong and they would go to the competitors with a better offering. Strategy was seen as aligning the whole organisation to the market position that could achieve long-term profitable differentiation when compared to competitors. Functional strategies were simply a more detailed interpretation of this overall imperative. Furthermore, strategy must be something that could be planned and directed. If managers could not influence strategy, then how could business be anything other than a lottery?
- The idea that sustainable competitive advantage could come from the capabilities of one's resources was a clear threat to the established position. Furthermore, the idea that strategies emerged, sometimes haphazardly and unpredictably over time, rather than were deliberate decisions taken by senior managers was also seemingly counterintuitive. Yet there is now considerable research evidence to support both these, once outrageous, propositions. However, widely practised approaches to developing operations strategies are still ignoring (or downplaying) the idea of resourced-based strategic advantage. For example, the business-model/operating-model idea is predominantly a 'top-down' philosophy that relegates operations capabilities to a 'supporting' rather than 'driving' role. The position we have taken in this chapter is one of blending some aspects of the traditional view with the more recent ideas. Nevertheless, it is important to understand that there are still different views on the very nature of strategic management.

SUMMARY CHECKLIST

- ☐ Does the operation have a fully articulated operations strategy?
- ☐ Does it include a vision for the role and contribution of the operations function?
- ☐ Does operations strategy take significant stakeholders into account?
- ☐ What position on the Hayes and Wheelwright Four-Stage Model are your operations?
- ☐ Is there a recognised process for translating business strategy 'top-down' into operations strategy?
- ☐ Does operations strategy demonstrate both correspondence and coherence with business strategy?

- ☐ Does the organisation's business model fit with its operating model?
- ☐ Are the operation's performance objectives fully articulated?
- ☐ Are performance objectives understood in terms of whether they are order-winners or qualifiers?
- ☐ Do different parts of the operation (probably producing different products or services) have their own relative priority of performance objectives that reflect their possibly different competitive positions?
- ☐ Is there a recognised process for bottom-up communication on strategic issues?
- ☐ Are the main strategic decisions that shape operations resources fully identified?
- ☐ Is the idea of operations-based capabilities fully understood?
- ☐ What capabilities does the operation currently possess?
- ☐ Are these operations and/or resources scarce, imperfectly mobile, imperfectly imitable or imperfectly substitutable?
- ☐ Are the logical links established between what the market requires (in terms of performance objectives) and what capabilities an operation possesses (in terms of the major strategic decision areas)?
- ☐ Where would you put the operation in terms of Figure 2.11, which describes the 'line of fit' between market requirements and operations capabilities?
- ☐ Have the key trade-offs for the operation been identified?
- ☐ What combination of repositioning, in order to change the nature of trade-offs and overcoming the trade-offs themselves, is going to be used to improve overall operations performance?

Case study

IKEA looks to the future[6]

For decades, IKEA has been one of the most successful retail operations in the world, with much of its success founded on how it organises its design, supply and retail service operations. With over 400 giant stores in 49 countries, IKEA has managed to develop its own standardised way of selling furniture. Its so-called 'big box' formula has driven IKEA to the global number one position in furniture retailing. 'Big box' because the traditional IKEA store is a vast blue-and-yellow maze of a showroom (on average around 25,000 square metres) where customers often spend around two hours – far longer than in rival furniture retailers. This is because of the way it organises its store operations. IKEA's philosophy goes back to the original business, started in the 1950s in Sweden by the late Ingvar Kamprad. He was selling furniture through a catalogue operation, and because customers wanted to see some of his furniture, he built a showroom on the outskirts of Stockholm and set the furniture out as it would be in a domestic setting. Also, instead of moving the furniture from the warehouse to the showroom area, he asked customers to pick the furniture up themselves from the warehouse. This approach became fundamental to IKEA's ethos – what has been called the 'we do our part you do yours' approach.

Ikea's 'big box' stores

Ikea offers a wide range of Scandinavian designs at affordable prices, usually stored and sold as a 'flat pack', which the customer assembles at home. '*It was an entirely new concept, and it drove the firm's success*', says Patrick O'Brien, Retail Research Director at retail consultancy GlobalData. '*But it wasn't just what IKEA was selling that was different, but how it was selling it.*' The stores were located and designed around one simple idea – that finding the store, parking, moving through the store itself, and ordering and picking up goods should be simple, smooth and problem-free. Catalogues are available at the entrance to each store showing product details and illustrations. For young children, there is a supervised children's play area, a small cinema, a parent and baby room and toilets, so parents can leave their children in the supervised play area for a time. Parents are recalled via the loudspeaker system if the child has any problems. Customers may also borrow pushchairs to keep their children with them.

Parts of the showroom are set out in 'room settings', while other parts show similar products together, so that customers can make comparisons. Given the volume of customers, there are relatively few staff in the stores. IKEA says it likes to allow customers to make up their own minds. If advice is needed, 'information points' have staff who can help. Every piece of furniture carries a ticket indicating its location in the warehouse from where it can be collected. Customers then pass into an area where smaller items are displayed that can be picked directly, after which they pass through the self-service warehouse where they can pick up the items they viewed in the showroom. Finally, customers pay at the checkouts, where a conveyor belt moves purchases up to the checkout staff. The exit area has service points and a large loading area allowing customers to bring their cars from the car park and load their purchases. Within the store a restaurant serves, among other things, IKEA's famous Swedish meatballs. IKEA's fans say they can make a visit to the store a real 'day out'.

But not everyone is a fan

Yet not all customers (even those who come back time after time) are entirely happy with the traditional IKEA retail experience. Complaints include:

- It can be a long drive to reach one of their stores (unless you are 'lucky' enough to live near one).
- The long 'maze-like' journey that customers are 'encouraged' to take through the store is too prescriptive.
- There are too few customer-facing staff in the store.
- There are long queues at some points in the store, especially at checkouts, and at busy times such as weekends.
- Customers have to locate, pick off the shelves and transport sometimes heavy products to the checkouts.
- IKEA designs can be 'bland' (or 'clean and aesthetically pleasing', depending on your taste).
- The furniture has to be assembled once you get it home, and the instructions are confusing.

Although many are

However, the impressive growth and success of IKEA over the years indicates that the company is doing many things right. Among the reasons customers give for shopping at IKEA are the following:

- Everything is available under one roof (albeit a very big roof).
- The range of furniture is far greater than at other stores.
- The products are displayed both by category (e.g. all chairs together) and in a room setting.
- Availability is immediate (competitors often quote several weeks for delivery).
- There is a kids' area and a restaurant so visiting the store is 'an event for all the family'.
- The design of furniture is 'modern, clean and inoffensive' – it fits anywhere.
- For the quality and design, the products are very good 'value for money'.

Was a new approach needed?

For decades, IKEA's unique retailing operations, combined with an excellent supply network and a customer-focused design philosophy, was an effective driver of healthy growth. However, there were indications that the company was starting to ask itself how it could solve some of the criticisms of its retail operations. '*We had to move away from conversations that began: "I love IKEA, but shopping at an* IKEA *store is not how I want to spend my time".*' (Gillian Drakeford, IKEA's UK boss) It needed to counter the complaints by some customers that its stores were understaffed, that the navigation of stores was too prescriptive and that queues were too long. '*We have had a great proposition for 60 years, but the customer had to fit around it, but the world has changed and to remain relevant we need to have a proposition that fits around the customer.*' (Gillian Drakeford)

IKEA was also realising that its 'big box' stores were under threat from a decline in car ownership (in 1994, 75 per cent of 21–29-year-olds held driving licences in the UK; by 2017 that had dropped to 66 per cent). Also, customers were increasingly wanting their flat-pack furniture to be delivered, rather than having to drive to an IKEA store to collect it. Ideally, they also wanted to order it online. IKEA did have an online presence, but compared to its competitors it was relatively underdeveloped. Not only that, but not all customers wanted to assemble their own furniture. '*The entire premise that IKEA developed was that consumers would be willing to drive their cars 50 kilometres to save some money on something that looks amazing*', said Ray Gaul, a retail analyst at Kantar Retail. '*Young people like IKEA, but they can't or don't want to drive to IKEA.*' However, the traditional 'big box' strategy was still popular with many customers, and sales from its stores continued to grow. Yet, in most markets, there were plenty of potential customers who could not reach an IKEA store within a reasonable drive (assumed to be around two-and-a-half hours). Some degree of rethinking IKEA's operating model seemed to be required. Torbjörn Lööf, CEO of Inter IKEA (who manage the IKEA concept), summarised the commitment to a rethink. '*We have been successful on a long journey. But it is clear that one era is ending and another beginning.*'

Smaller stores to complement the larger ones

Since 2015, IKEA has opened several smaller-footprint stores in Europe, Canada, China and Japan. But not all were the same. As a deliberate strategy, each was slightly different. This allowed the company to test alternative ways of locating, designing and managing its new ventures. Should they have cafés? How big should they be? Should they carry a range of products, or focus on a single category? Should they be located in shopping malls or on the high street? So, a 'pop-up' IKEA store in central Madrid offered only bedroom furnishings. A store in Stockholm focused on kitchens. It allowed customers to cook in the store, and book a 90-minute consultation to plan their kitchen. A small store in London stocked a range of product categories but had no café (only a coffee machine), and in place of a supervised kids' play area, computer games were provided. Other new stores were, in some ways, similar to traditional stores but smaller, with fewer car parking spaces, less inventory and acted as order-and-collection points. In some, customers could get expert advice on larger purchases, such as kitchens or bathrooms. Often in the smaller stores, only a few items could be purchased and taken home instantly. Rather, customers could use touch screens to order products and arrange for delivery or pick up at a convenient time. '*For me, it's a test lab for penetrating city centres*', said one senior IKEA executive. '*About 70 per cent of the people shopping there wouldn't go to a [traditional IKEA] store.*'

TaskRabbit

In 2017 IKEA bought TaskRabbit, whose app was one of the leaders in what was becoming known as the 'gig' economy. Using its app, over 60,000 independent workers or 'taskers' (at the time of acquisition) offered their services to customers wanting to hire someone to do tasks such as moving or assembling furniture. '*In a fast-changing retail environment, we continuously strive to develop new and improved products and services to make our customers' lives a little bit easier. Entering the on-demand, sharing economy enables us to support that*', IKEA Chief Jesper Brodin said in a statement. '*We will be able to learn from TaskRabbit's digital expertise, while also providing IKEA customers additional ways to access flexible and affordable service solutions to meet the needs of today's customer.*'

Web-based retailing

Arguably, the most significant retailing development in this period was the growth in online shopping. However, IKEA was slow to move online. This was partly because there was internal reluctance to interfere with its successful 'big box' retail operations, which encouraged customers to spend a long time in store, prone to impulse purchases. However, it became clear that the company needed to become fully committed to 'multichannel' retail operations, including online sales. It was also clear that there would not be a total shift to online sales. The idea was to offer both physical and digital options for customers who wanted to use both channels, and to win new customers online who would never make the journey to its 'big box' superstores. Some retail experts warned that the new strategy carried the same risks faced by all firms going online. According to Marc-André Kamel of consultants Bain & Company, '. . . *customers are not shifting entirely to e-commerce, but wish to mix and match channels*'. And, although IKEA had little choice but to invest in online channels, the danger was that it could raise costs, especially as the company was also planning significant bricks-and-mortar expansion in new markets such as India, South America and South-East Asia.

Third-party sales

Another break with traditional IKEA practice came when it announced that it would consider selling its products through independent 'third-party' online retailers. Torbjörn Lööf, CEO of Inter IKEA, said the decision to supply online retailers was an important part of the broader overhaul of its operations. '*[It] is the biggest development in how consumers meet Ikea since the concept was founded*', he told the *Financial Times*.

Sustainability

IKEA was among the world's biggest users of wood (estimated as around 1 per cent of all wood used), and some environmental groups condemned what they saw as the 'disposable' nature of its furniture. Responding to this criticism, IKEA appointed a Chief Sustainability Officer – the first time that sustainability was directly represented in the senior management team, and a recognition of the growing role of sustainability in determining how IKEA was perceived. It also recognised the ability for sustainability to drive business innovation, '. . . *we live in a world of finite resources and we recognise that consumption needs to reflect this. At IKEA, we are therefore seeking new ways to meet people's needs and aspirations whilst staying within the limits of our planet. . . We see emerging circular-economy business models as a great opportunity to develop the business further. If we didn't, we'd have to start looking at them as a risk. Sooner or later, other companies will start creating business models that disrupt the accepted way of selling home furnishings.'* (Jonas Engberg, Ikea Denmark's Sustainability Manager) In one initiative in Belgium, '. . . *we offer our customers five options to give furniture a second life: selling old IKEA-furniture in the store [at the price paid to the customer who supplied it], renewing it by repainting or reassembling, repairing by offering replacement parts, returning old furniture through our transport service, and donating to social organisations'*. Some commentators questioned the idea of selling longer-lasting products and trading pre-owned items without a mark-up as being bad for business. But Jonas Engberg disagreed: '*People sometimes come to IKEA with a bit of a guilty conscience. They want to buy stuff, but they are unable to completely forget the consequences. Our customers want this, they want us to transform the way we do business. When we started buying back our furniture at Aalborg we actually saw an increase in revenue.'*

Questions

1. In the traditional IKEA 'big box' stores, what is the relative importance of the operational performance objectives (quality, speed, dependability, flexibility, cost), compared to a conventional high-street furniture store?
2. What trade-offs are customers who go to these big stores making?
3. How does the strategy of increasing its online presence impact on these trade-offs?
4. An IKEA executive was reported as saying that in some parts of the world '. . . *we have reached the point of "peak stuff"*'. It was interpreted by some as a warning that consumer appetite for home furnishings had reached a crucial turning point. What are the implications of this for IKEA?

Applying the principles

Some of these exercises can be answered by reading the chapter. Others will require some general knowledge of business activity and some might require an element of investigation. **Model answers for the first two questions are available on the student companion website.** *For model answers to the other questions in this section, please ask your tutor/lecturer.*

1. The environmental services department of a city has two recycling services – newspaper collection (NC) and general recycling (GR). The NC service is a door-to-door collection service that, at a fixed time every week, collects old newspapers that householders have placed in reusable plastic bags at their gate. An empty bag is left for the householders to use for the next collection. The value of the newspapers collected is relatively small; the service is offered mainly for reasons of environmental responsibility. By contrast the GR service is more commercial. Companies and private individuals can request a collection of materials to be disposed of, either using the telephone or the internet. The GR service guarantees to collect the material within 24 hours unless the customer prefers to specify a more convenient time. Any kind of material can be collected and a charge is made depending on the volume of material. This service makes a small profit because the revenue both from customer charges and from some of the more valuable recycled materials exceeds the operation's running costs. How would you describe the differences between the performance objectives of the two services?

2. The Managing Partner of The Brandfair Partnership (TBP) describes her business: '*It is about four years now since we specialised in the small-to-medium firms' market. Before that we also used to provide brand consultancy services for anyone who walked in the door. So now we have built up our brand consultancy skills in many areas. However, within the firm, I think we could focus our activities even more. There seem to be two types of assignment that we are given. About 40 per cent of our work is relatively routine. Typically, these assignments are conventional market research and focus group exercises. Both these activities involve a standard set of steps that can be carried out by comparatively junior staff. Of course, an experienced consultant is needed to make some decisions; however, most of this work is fairly straightforward. Customers expect us to be relatively inexpensive and fast in delivering the service. They do not expect us to make simple errors; in fact if we did this too often we would lose business. Fortunately, our customers know that they are buying a "standard package" and don't expect it to be too customised. The problem here is that specialist agencies have been emerging over the last few years and they are starting to undercut us on price. Yet, I still feel that we can operate profitably in this market and anyway, we still need these capabilities to serve our other clients. The other 60 per cent of our work is for clients who require far more specialist services, such as assignments involving major brand reshaping. These assignments are complex, large, take longer, and require significant branding skill and judgement. It is vital that clients respect and trust the advice we give them in all "brand associated" areas such as product development, promotion, pricing, and so on. Of course, they assume that we will not be slow or unreliable in preparing advice, but mainly it's trust in our judgement backed up by hard statistics that is important to the client. This is popular work with our staff. It is both interesting and very profitable*.' How different are the two types of business described by the Managing Partner of TBP? It has been proposed that she split the firm into two separate businesses; one to deal with routine services and the other to deal with more complex services. What would be the advantages and disadvantages of doing this?

3. DSD designs, makes and supplies medical equipment to hospitals and clinics. Its success was based on its research and development culture. Although around 50 per cent of manufacturing was done in-house, its products were relatively highly priced, but customers were willing to pay for its technical excellence and willingness to customise equipment. Around 70 per cent of all orders involved some form of customisation from standard 'base models'. Manufacturing could take three months from receiving the specification to completing assembly, but customers were more interested in equipment being delivered on time rather than immediate availability. According to its CEO, '*manufacturing is really a large laboratory. The laboratory-like culture helps us to maintain our superiority in leading-edge product technology and customisation. It also means that we can call upon our technicians to pull*

out all the stops in order to maintain delivery promises. However, I'm not sure how manufacturing, or indeed the rest of the company, will deal with the new markets and products which we are getting into.'

The new products were 'small black box' products that the company had developed. These were devices that could be attached to patients, or implanted. They took advantage of sophisticated electronics and could be promoted directly to consumers as well as to hospitals and clinics. The CEO knew their significance. '*Although expensive, we have to persuade healthcare and insurance companies to encourage these new devices. More problematic is our ability to cope with these new products and new markets. We are moving towards being a consumer company, making and delivering a higher volume of more standardised products where the underlying technology is changing fast. We must become faster in our product development. Also, for the first time, we need some kind of logistics capability. I'm not sure whether we should deliver products ourselves or subcontract this. Manufacturing faces a similar dilemma. On one hand it is important to maintain control over production to ensure high quality and reliability; on the other hand, investing in the process technology to make the products will be very expensive. There are subcontractors who could manufacture the products; they have experience in this kind of manufacturing but not in maintaining the levels of quality we will require. We will also have to develop a "demand fulfilment" capability to deliver products at short notice. It is unlikely that customers would be willing to wait the three months our current customers tolerate. Nor are we sure of how demand might grow. I'm confident that growth will be fast but we will have to have sufficient capacity in place not to disappoint our new customers. We must develop a clear understanding of the new capabilities that we will have to develop if we are to take advantage of this wonderful market opportunity.*'

What advice would you give DSD? Consider the operational implications of entering this new market.

4. Xexon7 is a specialist artificial intelligence (AI) development firm that develops algorithms for various online services. As part of its client services it has a small (ten-person) helpdesk call centre to answer client queries. Clients can contact them from anywhere in the world at any time of the day or night with a query. Demand at any point in time is fairly predictable, especially during the (European) daytime. Demand during the night hours (Asia and the Americas) is considerably lower than in the daytime and also less predictable. '*Most of the time we forecast demand pretty accurately, so we can schedule the correct number of employees to staff the work stations. There is still some risk, of course. Scheduling too many staff at any point in time will waste money and increase our costs, while scheduling too few will reduce the quality and response of the service we give*.' (Peter Fisher, Helpdesk Manager) Peter was, overall, pleased with the way in which his operation worked. However, he felt that a more systematic approach could be taken to identifying improvement opportunities: '*I need to develop a logical approach to identify how we can invest in improving things like sophisticated diagnostic systems. We need to both reduce our operating costs and maintain, and even improve, our customer service.*' What are the trade-offs that must be managed in this type of call centre?

5. Visit the website of any large cement and aggregate company. How does it present its CSR strategy? What do you think are the difficulties of forming a CSR strategy in this type of company?

Notes on chapter

1 The information on which this example is based is taken from: company website, www.sstl.co.uk [accessed 10 September 2020]; Pfeifer, S. (2018) 'Brexit restrictions force UK satellite maker to send work to Europe', *Financial Times*, 19 November.

2 The information on which this example is based is taken from: B Corp website, https://bcorporation.net/about-b-corps [accessed 10 September 2020]; Economist (2018) 'Choosing plan B – Danone rethinks the idea of the firm', Business section, *Economist* print edition, 9 August.

3 Based on an example from Slack, N. (2017) *The Operations Advantage*: *A practical guide to making operations work*, Kogan Page. Used by permission.

4 See Skinner, W.C. (1969) 'Manufacturing – missing link in corporate strategy', *Harvard Business Review*, 1 May.

5 The information on which this example is based is taken from: Eley, J. (2020) 'How Dave Lewis fixed ailing supermarket giant Tesco', *Financial Times*, 30 March; Vandevelde, M. (2016) 'Tesco ditches global ambitions with retreat to UK', *Financial Times*, 21 June; Clark, A. and Ralph, A. (2014) 'Tesco boss defiant amid 4% plunge in sales', *The Times*, 5 June.

6 The information on which this case is based is taken from: IKEA website, https://www.ikea.com/gb/en/this-is-ikea/about-us/ [accessed 29 September 2020]; Matlack, C. (2018) 'The tiny Ikea of the future, without meatballs or showroom mazes', *Bloomberg Businessweek*, 10 January; Milne, R. (2018) 'What will Ikea build next?', *Financial Times*, 1 February; Economist (2017) 'Frictionless furnishing: IKEA undertakes some home improvements', *Economist* print edition, 2 November; Hipwell, D. (2017) "This is no time to sit back and relax – we must deliver", says IKEA's UK boss', *The Times*, 10 February; Gerschel-Clarke, A.T. (2016) '"Peak stuff": Why IKEA is shifting towards new business models', Sustainablebrands.com, 17 February [accessed 10 September 2020]; Milne, R. (2017) 'Ikea turns to ecommerce sites in online sales push', *Financial Times*, 9 October; Hope, K. (2017) 'Ikea: Why we have a love – hate relationship with the Swedish retailer', BBC News, 17 October; Armstrong, A. (2017) 'Revealed: How after 30 years, Ikea is undergoing a radical overhaul', *The Telegraph*, 15 October.

Taking it further

Barney, J.B. and Clark, D.N. ***(2007) Resource-Based Theory: Creating and sustaining competitive advantage, Oxford University Press.*** *Great on the 'resource-based view' of the firm.*

Braithwaite, A. and Christopher, M. ***(2015) Business Operations Models: Becoming a disruptive competitor, Kogan Page.*** *Aimed at practitioners, but authoritative and interesting.*

Gray, D., Micheli, P. and Pavlov, A. ***(2015) Measurement Madness: Recognizing and avoiding the pitfalls of performance measurement, John Wiley & Sons.*** *Lots of examples of how companies can misuse performance measurement.*

Hayes, R. ***(2006) 'Operations, strategy, and technology: pursuing the competitive edge', Strategic Direction, 22 (7).*** *A summary of the subject from one (if not the) leading academics in the area.*

Hayes, R.H. and Pisano, G.P. ***(1994) 'Beyond world class: the new manufacturing strategy', Harvard Business Review, 72 (1).*** *Same as above.*

Hill, A. and Hill, T. ***(2009) Manufacturing Operations Strategy, 3rd edition, Palgrave Macmillan.*** *The descendant of the first non-US book to have a real impact in the area.*

Johnson, G., Whittington, R. and Scholes, K. ***(2011) Exploring Strategy, 9th edition, Pearson.*** *There are many good books on strategy. This is one of the best.*

Savitz, A.W. and Weber, K. ***(2006) The Triple Bottom Line: How today's best-run companies are achieving economic, social and environmental success – and how you can too, Jossey-Bass.*** *Good on the triple bottom line.*

Slack, N. and Lewis, M. ***(2020) Operations Strategy, 6th edition, Pearson.*** *What can we say – just brilliant!*

Slack, N. ***(2017) The Operations Advantage: A practical guide to making operations work, Kogan Page.*** *Apologies for self-referencing again! This book is written specifically for practitioners wanting to improve their own operations – short and to the point!*

3 Product and service innovation

Introduction

Customers want innovation, and the businesses that provide it can succeed to a greater degree than their competitors that are content to offer the same things (or services) to their customers, created in the same old way. Many successful companies are successful because they have refreshed the idea of what their markets want. Their products and services have been continually updated, altered and modified. Some changes are small, incremental adaptations to existing ways of doing things. Others are radical, major departures from anything that has gone before. The innovation activity is about successfully delivering change in its many different forms. Being good at innovation has always been important. What has changed in recent years is the sheer speed and scale of innovation in industries all over the world. Innovation processes are also becoming more complex, with inputs from different individuals and departments within an organisation, and increasingly from a wide variety of external sources. Operations managers are expected to take a greater and more active part in service and product innovation. This chapter examines what is meant by innovation and why it matters; the general stages involved in bringing service and product offerings from concept to launch; and some of the resourcing and organisational considerations for innovation (see Figure 3.1).

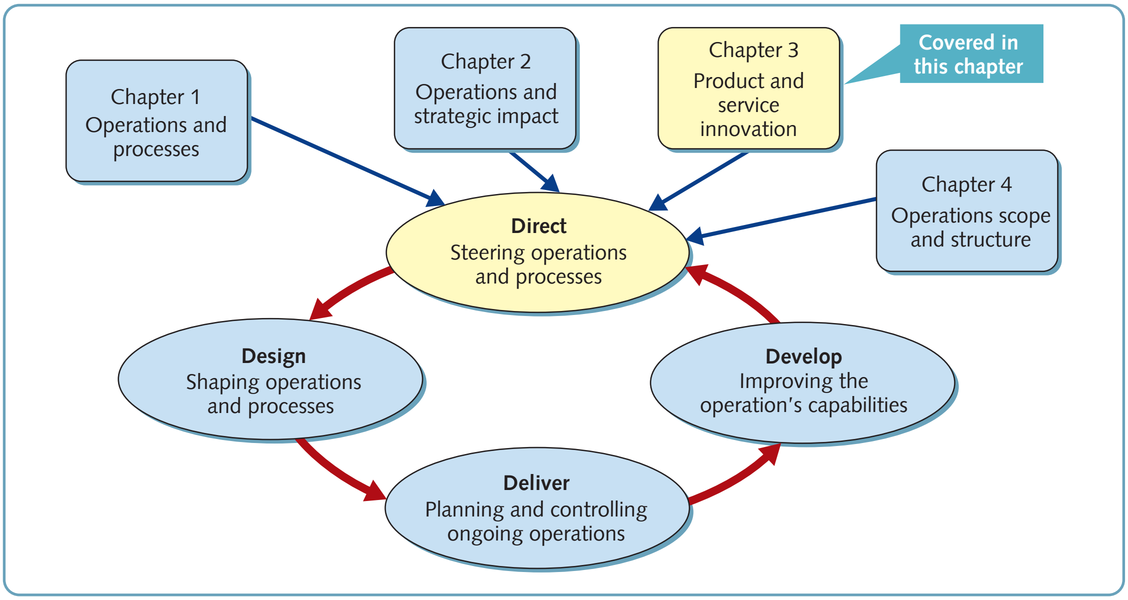


Figure 3.1 Product and service innovation

EXECUTIVE SUMMARY

3.1 What is the strategic role of product and service innovation?

↓

3.2 Are the product and service innovation process objectives specified?

↓

3.3 Is the product and service innovation process defined?

↓

3.4 Are the resources for developing product and service innovation adequate?

↓

3.5 Is the development of products and services and of the process that created them simultaneous?

3.1 What is the strategic role of product and service innovation?

A business's products and services are how markets judge it: they are its 'public face'. Innovation is all about doing things differently. That is why product and service innovation is so important. It involves changing what is offered to customers in some way, in addition to using 'creativity' and 'design' – terms that have similar but different meanings. Creativity is the use of imagination or original ideas; it is an essential ingredient of innovation, which means doing something new. The design activity transforms innovation into a practical proposition by defining its looks or arrangement. Innovation often follows an S-shaped progress, where, in the early stages of the introduction of new ideas, relatively small performance improvements are experienced, but, with time, performance increases and eventually levels off. A further strategic perspective on innovation is the difference between incremental and radical innovation. It is a distinction, together with that between 'architectural' and 'component' knowledge in the Henderson–Clark model. Yet, at its most basic, the innovation activity is a process that involves many of the same design issues common to other operations processes.

3.2 Are the product and service innovation process objectives specified?

The performance of the innovation process can be assessed in much the same way as we would consider the outputs from it. The quality of the innovation process can be judged in terms of both conformance (no errors in the offering) and specification (the effectiveness of the offering in achieving its market requirements). Speed in the innovation process is often called 'time to market' (TTM). Short TTM also implies that the offering can be introduced to the market frequently, achieving strategic impact. Dependability in the innovation process means meeting launch delivery dates. In turn, this often requires that the innovation process is sufficiently flexible to cope with disruptions. The cost of the innovation process can be thought of both as the amount of budget that is necessary to develop a new offering and as the impact of the innovation on the cost of delivering the service or product on an ongoing basis. Finally, the sustainability of the innovation process considers the impact on broad stakeholder objectives encompassed within the 'triple bottom line' – people, planet and profit.

3.3 Is the product and service innovation process defined?

To create a fully specified service or product offering, potential designs tend to pass through a set of stages in the innovation process. Almost all stage models start with a general idea or 'concept' and progress through to a fully defined specification for the offering, incorporating various service and product components. In-between these two states, the offering may pass through the following stages: concept generation, concept screening, preliminary design (including consideration of standardisation, commonality, modularisation and mass customisation), evaluation and improvement, prototyping and final design.

3.4 Are the resources for developing product and service innovation adequate?

To be effective, the innovation process needs to be resourced adequately. The detailed principles of process design that are discussed in Chapters 5 and 6 are clearly applicable when developing new services or product offerings. However, because the innovation activity is often an operation in its own right, there are some more strategic issues to consider: how much capacity to devote to innovation; how much of the innovation activity to outsource; and what kinds of technology to use in the development of new offerings.

3.5 Is the development of products and services and of the process that created them simultaneous?

The outputs from the innovation, in the form of new service and product offerings, are important inputs into the processes that create and deliver them on an ongoing basis. Therefore, it is often best to consider these in parallel rather than (as more traditionally done) in sequence. Merging the innovation process for new service and product offerings with the processes that create them is sometimes called simultaneous (or interactive) design. Its key benefit is the reduction in the time taken for the whole innovation activity. In particular, four simultaneous design factors can be identified that promote fast time to market. These are: routinely integrating the design of the product–service offering and the design of the process used to create and deliver them; overlapping the stages in the innovation process; the early deployment of strategic decision-making to resolve design conflict; and an organisational structure that reflects the nature of the offering.

3.1 Diagnostic question: *What is the strategic role of product and service innovation?*

A business's products and services are its 'public face'. They are what markets judge a company on: good products and services equals a good company. This is why it is important to devote time and effort to the way in which innovative ideas are incorporated in new products and services. Moreover, it has long been accepted that there is a connection between how companies go about developing innovative products and services and how successful those products and services are in the market-place.

Innovation, design and creativity

Developing new products and services is a creative, and often innovative, process. What is the relationship between terms such as 'innovation', 'creativity' and 'design'? They have similar but different meanings, overlap to some extent and are clearly related to each other. There are many definitions of 'innovation', and the term is notoriously ambiguous and lacks either a single definition or measure. Here, we use the word to mean a new idea or knowledge that leads to some new aspect of performance, with a stress on novelty and change. In other words, innovation is simply about doing something new. However, the idea of innovation is both broader and more complete than that of 'invention'. An 'invention' is also something that is novel or unique (usually applied to a device or method), but it does not necessarily imply that the novel device or method has the potential to be practical, economic or capable of being developed commercially. Innovation goes further, it implies that the novel idea can provide a return for an organisation's customers, owners or both. The study of innovation, including what influences it and how to manage it, is a huge subject. However, creativity is one particularly important attribute that is central to innovation. 'Creativity' is the ability to move beyond conventional ideas, rules or assumptions, in order to generate significant new ideas. It is a vital ingredient in innovation.

If creativity is an essential ingredient of innovation, and innovation implies making novel ideas into practical, commercial form, the process that transforms innovative ideas into something more concrete is 'design'. Innovation creates the novel idea; design makes it work in practice. Design is to conceive the looks, arrangement, and workings of something. A design must deliver a solution that will work in practice. Design is also an activity that can be approached at different levels of detail. One may envisage the general shape and intention of something before getting down to defining its details. Figure 3.2 illustrates the relationship between creativity, innovation and design as we use the terms here. These concepts are intimately related, which is why we treat them in the same chapter. First, we will look at some of the basic ideas that help to understand innovation.

OPERATIONS PRINCIPLE
Creativity and innovation both contribute to the effectiveness of the design of a product or service.

Innovation is all about doing things differently. However, not all innovation implies fast adoption, even when an idea has merit. Look at the history of the zip in the case example 'The slow innovation progress of the zip fastener'. The zip was not the result of any radical scientific discovery. It was simply an ingenious

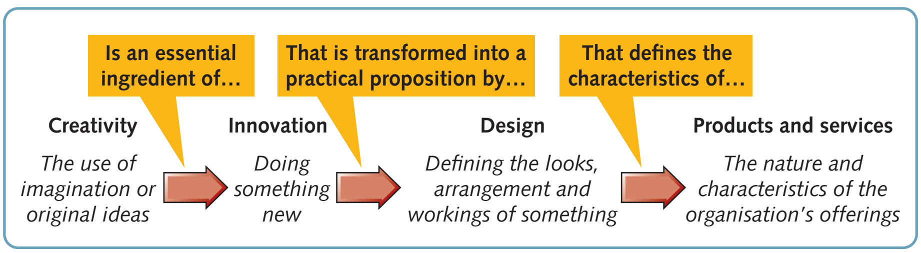


Figure 3.2 **The relationship between creativity, innovation and design**

Source: From Slack, N. *Operations Management*, 8e, © 2016 Pearson Education Limited, UK.

Case example

The slow innovation progress of the zip fastener[1]

Some innovations take a long time to become successful. The zip (zipper, in the US) is one such innovation. In 2017, the market for zip fasteners was around $11 billion, and is expected to grow to around $20 billion by 2024. This is based on the increasing global demand for clothing, luggage and other products that use zips, alongside the acceleration of fast fashion. Historically, it is a relatively recent innovation. For centuries, clothes were held together with loops and toggles, buckles, brooches, laces, or simply tied and wrapped. It was in 14th-century Britain that the hook and eye, what could be considered the earliest ancestor of the zip, began to be used. But they were both awkward to use and fragile. The first real zip-like device was patented in the US by Whitcomb Judson's Universal Fastener Company of Chicago in 1893. His design specified a sliding guide to pull together a line of hooks and a line of eyes on a boot. Unfortunately, it proved rather unreliable and the company was taken over by Gideon Sundback, an engineer from Sweden. His innovation abandoned hooks and eyes and replaced them with rows of metal protuberances with a tooth on one side and a socket on the other, similar in principle to today's design. Around the same time, a Swiss woman called Katharina Kuhn-Moos patented a similar design, but it was never manufactured. But the device was still expensive compared to more conventional buttons.

It was the Japanese company YKK that transformed the zip's prospects. Although there had been developments of the basic design, for example, plastic teeth to replace metal ones and continuous spirals of nylon used as teeth, the real breakthrough came from improving the quality in how zips were manufactured. Accuracy of manufacturing is important in how well a zip works. Very small misalignments can cause jams and breakages and, unlike losing a button, a broken zip can mean discarding a garment. Tadao Yoshida, known as 'the zipper king', had founded YKK in 1934, and had gained such a reputation for quality and reliability (guarantees that each of its zips will last for 10,000 uses) that in 1960, when Sundback's patents expired, YKK was able to move into the larger US market. It went on to gain around 40 per cent of the market, by value, and makes more zips every year than there are people on the planet.

way of making life a bit easier, and it happened to become popular (eventually), whereas there was clearly a need for Edison's light-bulb filament, which made mass-illumination possible. Similarly, Bell's telephone made possible something that had been previously impossible. Yet, with the zip, there was no consensus that fastenings was an area in dire need of innovation.

The innovation S-curve

When new ideas are introduced in services, products or processes, they rarely have an impact that increases uniformly over time. Usually performance follows an S-shaped progress. So, in the early stages of the introduction of new ideas, although often large amounts of resources, time and effort are needed to introduce the idea, relatively small performance improvements are experienced. However, with time, as experience and knowledge about the new idea grow, performance increases. Nevertheless, as the idea becomes established, extending its performance further becomes increasingly difficult – see Figure 3.3(a). When one idea reaches its mature, 'levelling-off' period, it is vulnerable to a further new idea being introduced that, in turn, moves through its own S-shaped progress. This is how innovation works: the limits of one idea being reached prompts a newer, better idea, with each new S-curve requiring some degree of redesign – see Figure 3.3(b).

> OPERATIONS PRINCIPLE
> *Usually the performance of each new innovation follows an S-shaped progress.*

Incremental or radical innovation

An obvious difference between how the pattern of new ideas emerges in different operations or industries is the rate and scale of innovation. Some industries, such as telecommunications, enjoy frequent and often significant 'breakthrough' or 'radical' innovations. Others, such as house building, do have innovations, but they are usually less dramatic and more 'continuous'

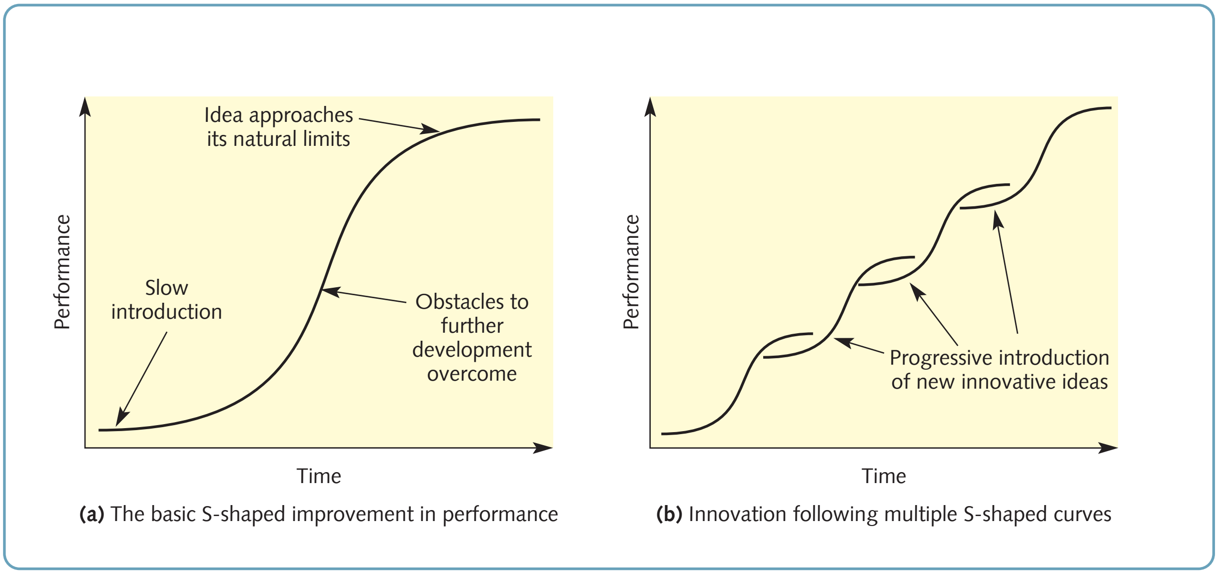


Figure 3.3 The S-shaped curve of innovation

or 'incremental'. Radical innovation often includes large technological advancements that may require completely new knowledge and/or resources, making existing services and products obsolete and therefore non-competitive. Incremental innovation, by contrast, is more likely to involve relatively modest technological changes, building upon existing knowledge and/or resources. This is why established companies may favour incremental innovation because they have built up a significant pool of knowledge (on which incremental innovation is based). In addition, established companies are more likely to have a mindset that emphasises continuity, perhaps not even recognising potential innovative opportunities. New entrants to markets, however, have no established position to lose, nor do they have a vast pool of experience. They may be more likely to try for innovation that is more radical.

The Henderson–Clark Model

Although distinguishing between incremental and radical innovation is useful, it does not fully make clear why some companies succeed or fail at innovation. Two researchers, Henderson and Clark, looked at the question of why some established companies sometimes fail to exploit seemingly obvious incremental innovations. They answered this question by dividing the technological knowledge required to develop new services and products into 'knowledge of the components of knowledge' and 'knowledge of how the components of knowledge link together'. They called this latter knowledge 'architectural knowledge'. Figure 3.4 shows what has become known as the Henderson–Clark Model.[2] It refines the simpler idea of the split between incremental and radical innovation. In this model, incremental innovation is built upon existing component and architectural knowledge, whereas radical innovation changes both component and architectural knowledge. Modular innovation builds on existing architectural knowledge, but requires new knowledge for one or more components. By contrast, architectural innovation will have a great impact upon the linkage of components (or the architecture), but the knowledge of single components is unchanged.

For example, in healthcare services, simple (but useful and possibly novel at the time) innovations in a primary healthcare (general practitioner) doctors' clinic, such as online appointment websites, would be classed as incremental innovation because neither any elements nor the relationship between them is changed. If the practice invests in a new diagnostic heart scanner, that element of their diagnosis task has been changed and will probably need new knowledge,

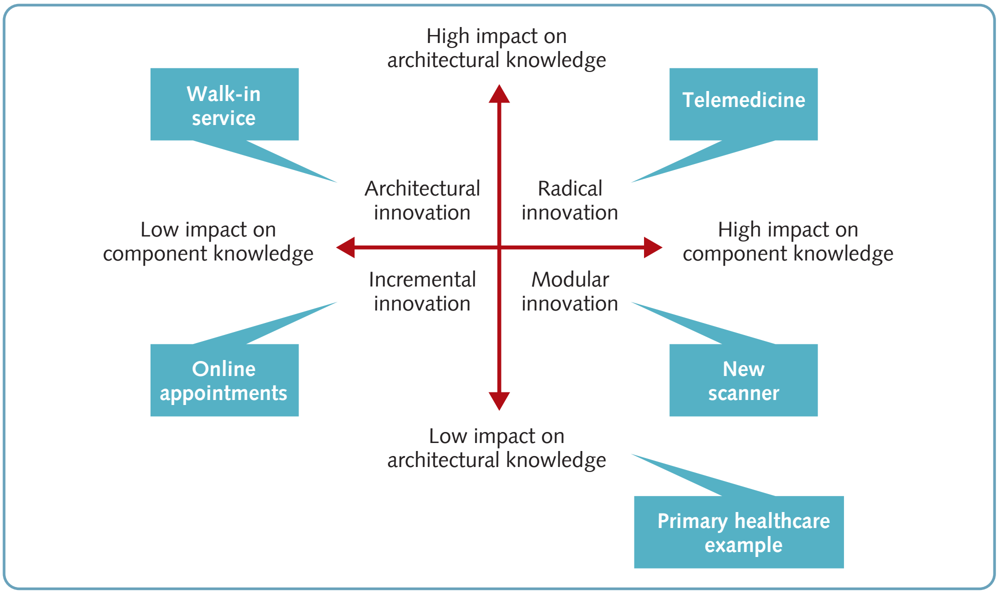


Figure 3.4 **The Henderson–Clark Model**

> **OPERATIONS PRINCIPLE**
> ***Innovation can be classified on two dimensions: innovation of components of a design and innovation of the linkages between them.***

but the overall architecture of the service has not been changed. This innovation would be classed as 'modular'. An example of architectural innovation would be the practice providing 'walk-in' facilities in the local city centre. It would provide more or less the same service as the regular clinic (no new components), but the relationship between the service and patients has changed. Finally, if the practice adopted some of the 'telemedicine' technology, that would represent radical innovation.

Innovation is influenced by later stages in the value chain

Many illustrations of how an enterprise creates and captures value start with the innovation/design stage. At its simplest, a firm innovates in the form of a design for a product or service (or some blend of the two), produces or creates it through its core operations and distributes it to its customers, who then use or experience it. Each of these stages is a transformation process. Innovation/design transforms ideas into workable designs. Production/creation transforms the design into a form that customers will find useful. Distribution disseminates it (physically or virtually). Finally, customers gain value by using it. But these stages are not independent of each other. Certainly, the innovation/design process is influenced by all the subsequent stages. Most of us are used to thinking about how product or service designs are judged primarily by how they add value for users, but both production and distribution stages can also impact the design stage (see Figure 3.5).

Design for production

Decisions taken during the design of a product or service can have a profound effect on how they can be created. For physical products, this has been well understood for decades and is usually termed 'design for production' (DFP), 'design for manufacture' (DFM) or 'design for manufacture and assembly' (DFMA). But the same principle applies equally to services. How a service is designed and specified can make its execution easy or difficult in practice. The design of queuing areas in a theme park attraction could entertain waiting customers or irritate them. The scripts used by call-centre staff can be written to be appropriate to a wide range of

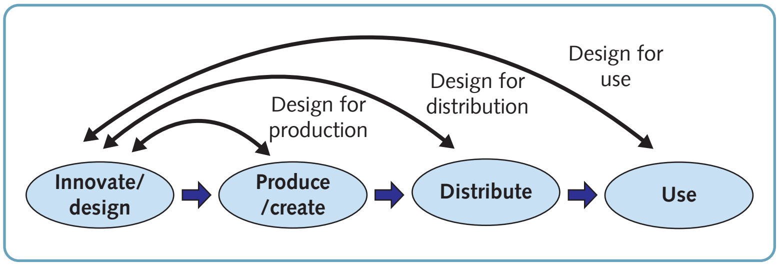


Figure 3.5 The influences on the design of a product or service are not limited to how it will be used, but include how it will be created and distributed

enquiries, or sound confusing in some situations, and so on. Technology can help. Virtual reality can help service engineers to (virtually) 'walk round' facilities so that maintenance procedures can be designed. Designers, or customers, can be (virtually) inside sports venues, aeroplanes, buildings and amusement parks, for example.

Design for distribution

The most obvious example of how the design of a product can be influenced by how it is distributed is 'flat pack' furniture. IKEA stores exemplify the extent to which clever 'flat pack' design can influence subsequent stages in the value chain. Designing IKEA furniture to be sold in flat form allows for efficient transportation and efficient storage, which in turn allows customers in its stores to collect it. Similarly, some products are designed so that, when packaged, they fit conveniently onto transportation pallets or containers. Again, the same idea applies to services. The design of online services can be influenced by how the service could be presented through its web page. Even purely 'artistic' offerings such as music are influenced by the way they are distributed. For example, the majority of revenue in the music industry comes from 'product' distributed through streaming services, where artists are paid per play – provided, that is, that the listener plays the piece for at least 30 seconds. About one-third of all streams are played because a track has been included on a streaming company's playlist, usually selected by algorithms, the precise form of which are not always known. It is often assumed that the algorithms favour tracks that get to the 'catchy' part of the tune relatively fast. Some point out that this has led to the production of music tracks that are shorter, with truncated introductions and choruses that start sooner.[3]

Product and service design innovation as a process

Although differing slightly, the innovation process is essentially similar across a whole range of industries. More importantly, it is a process. The aim of the innovation process is first to create offerings that exceed customers' expectations in terms of quality, speed, dependability, flexibility, cost and sustainability; and second, to ensure that competitors find these offerings hard to imitate, substitute or gain access to. With increasingly demanding customers, higher levels of competition and shorter product–service life cycles in many markets, organisations that can master the process of innovation may gain significant competitive advantage. And, like any other process, the innovation process should be managed as a process. Figure 3.6 uses the input–transformation–output model to describe the innovation process. Inputs come both from within the organisation – employees, R&D, operations, marketing, human resources and finance, and from outside it – market research, customers, lead users, suppliers, competitors, collaborators and wider stakeholders. The outputs (or outcomes) of the innovation process include not just the details of the offering, but also an understanding of its value proposition.

> OPERATIONS PRINCIPLE
> ***The innovation activity is a process that can be managed using the same principles as other processes.***

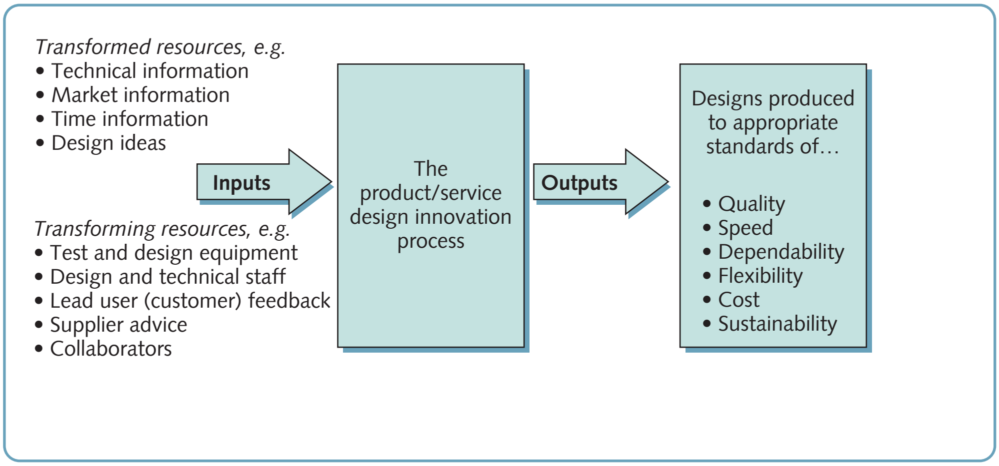


Figure 3.6 The product or service design innovation activity as a process

3.2 Diagnostic question: *Are the product and service innovation process objectives specified?*

> **OPERATIONS PRINCIPLE**
> ***Innovation processes can be judged in terms of their levels of quality, speed, dependability, flexibility, cost and sustainability.***

The performance of the innovation process can be assessed in much the same way as we would consider the products and services that result from it – namely in terms of quality, speed, dependability, flexibility, cost and sustainability. These performance objectives have just as much relevance for innovation as they do for the ongoing delivery of offerings once they are introduced to the market.

What is the quality of the innovation process?

Design quality is not always easy to define precisely, especially if customers are relatively satisfied with existing service and product offerings. Many software companies talk about the 'I don't know what I want, but I'll know when I see it' syndrome, meaning that only when customers use the software are they in a position to articulate what they do or don't require. Nevertheless, it is possible to distinguish high- and low-quality designs (although this is easier to do in hindsight) by judging them in terms of their ability to meet market requirements. In doing this, the distinction between the specification quality and the conformance quality of designs is important. No business would want a design process that was indifferent to the creation of 'errors' in its designs, yet some are more tolerant than others. For example, in pharmaceutical development the potential for harm is particularly high because drugs directly affect our health. This is why the authorities insist on such a prolonged and thorough design process. Far more frequent are the 'product recalls' that are relatively common in, say, the automotive industry. Many of these are design related and the result of 'conformance' failures in the design process. The 'specification' quality of design is different. It means the degree of functionality, or experience, or aesthetics, or whatever the product or service is primarily competing on. Some businesses require product or service designs that are relatively basic (although free from errors), while others require designs that are clearly special in terms of the customer response they hope to elicit.

What is the speed of the innovation process?

The speed of innovation matters more to some industries than others. For example, innovation in construction and aerospace happens at a much slower pace than in clothing or microelectronics. However, rapid innovation or 'time-based competition' has become the norm for an increasing number of industries. Sometimes this is the result of fast-changing consumer fashion. Sometimes a rapidly changing technology base forces it. Yet, no matter what the motivation, fast design brings a number of advantages:

- *Early market launch* – increasing the speed of services or product development allows earlier marketing of new offerings that may command price premiums and generate revenues for longer.
- *Starting design late* – may have advantages, especially where either the nature of customer demand or the availability of technology is uncertain, so fast design allows design decisions to be made closer to the time when service and product offerings are introduced to the market.
- *Frequent market stimulation* – rapid innovations allow frequent new or updated offerings to be introduced.

What is the dependability of the innovation process?

Any benefits of fast design processes can be negated unless designs are delivered dependably. Design schedule slippage can extend design times, but also adds to the uncertainty surrounding the innovation process. Suppliers who do not deliver solutions on time, customers or markets that change during the innovation process, and so on, all contribute to an uncertain and ambiguous design environment. Professional project management (see Chapter 15) can help to reduce uncertainty and prevent (or give early warning of) missed deadlines, process bottlenecks and resource shortages. However, external disturbances to the innovation process will remain. These may be minimised through close liaison with suppliers and market or environmental monitoring. Yet, unexpected disruptions will always occur and the more innovative the design, the more likely they are to occur. This is why flexibility within the innovation process can help with dependable delivery of new service and product offerings.

What is the flexibility of the innovation process?

Flexibility in the innovation process is the ability to cope with external or internal change. The most common reason for external change is that markets, or specific customers, change their requirements. Although there may be no need for flexibility in relatively predictable markets, it is valuable in more fast-moving and volatile markets, where one's own customers and markets change, or where the designs of competitors' offerings dictate a matching or leapfrogging move. The increasing complexity and interconnectedness of service and product components in an offering may also require flexibility. A bank, for example, may bundle together a number of separate services for one particular segment of its market. Changing one aspect of this package may require changes in other elements. One way of measuring innovation flexibility is to compare the cost of modifying a design in response to such changes against the consequences to profitability if no changes are made. The lower the cost of modifying an offering in response to a given change, the higher is the level of flexibility.

What is the cost of the innovation process?

The cost of innovation is usually analysed in a similar way to the ongoing cost of delivering offerings to customers. These cost factors are split up into three categories: the cost of buying the inputs to the process; the cost of providing the labour in the process; and the other general

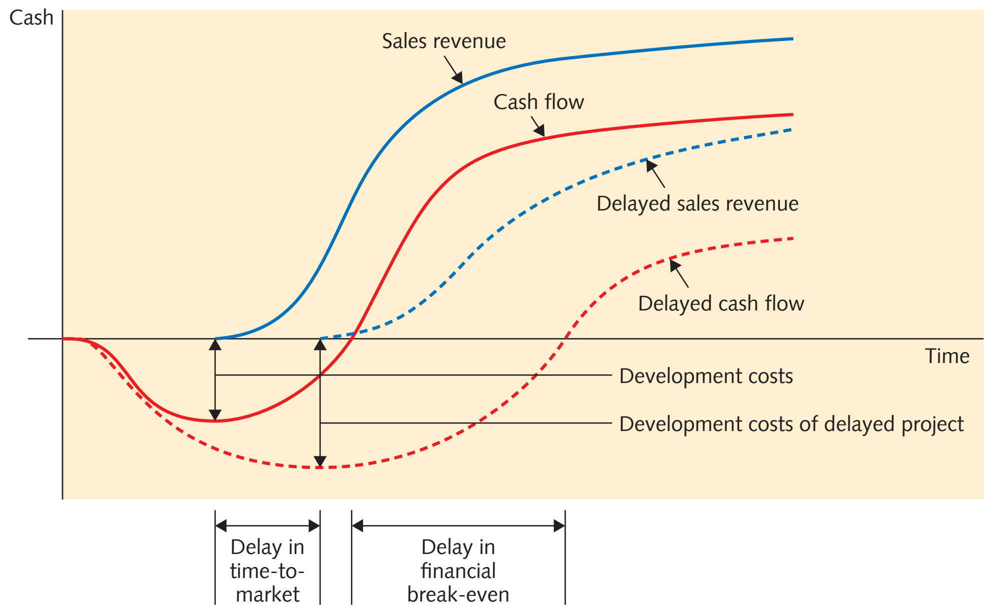


Figure 3.7 Delay in time-to-market of new products and services not only reduces and delays revenues, it also increases the costs of development. The combination of both of these effects usually delays the financial break-even point far more than the delay in the time to market

overhead costs of running the process. In most in-house innovation processes, the last two costs outweigh the first. But all aspects of performance can affect costs. So, whether caused by quality errors, a slow innovation process, a lack of project dependability, or delays caused through inflexibility, the end result is that the design will be late. This is likely to mean more expenditure on the design and delayed (and probably reduced) revenue. The combination of these effects usually means that the financial break-even point for a new offering is delayed far more than the original delay in its launch (see Figure 3.7).

What is the sustainability of the innovation process?

The product and service innovation process is particularly important in ultimately impacting on the ethical, environmental and economic well-being of stakeholders. Some innovation activity focuses particularly on the ethical dimension of sustainability. Banks have moved to offer ethical investments that seek to maximise social benefit as well as financial returns (avoiding businesses involved in weaponry, gambling, alcohol and tobacco, for example). Other examples of ethically focused innovations include the development of 'fair-trade' products such as tea, coffee, chocolate, cotton and handicrafts; clothing manufacturers establishing ethical trading initiatives with suppliers; supermarkets ensuring animal welfare for meat and dairy, and paying fair prices for vegetables; and online companies establishing customer complaint charters. Innovation may also focus on changing materials in the design to reduce its environmental burden. Examples include the use of organic cotton or bamboo in clothing; wood or paper from managed forests used in garden furniture, stationery and flooring; and recycled materials and natural dyes in clothing, curtains and upholstery. Other innovations may focus on how products and services are used. Computer designers may introduce power management systems. In the detergent industry, companies have developed products that allow clothes to be washed at lower temperatures. Architects can design houses that operate with minimal energy, or use

Case example

Product innovation for the circular economy[4]

Design innovation is not just confined to the initial conception of a product; it also applies to the end of its life. This idea is often called 'designing for the circular economy'. The 'circular economy' is proposed as an alternative to the traditional linear economy (or make–use–dispose as it is termed). The idea is to keep products in use for as long as possible, extract the maximum value from them while in use, and then recover and regenerate products and materials at the end of their service life. The circular economy is, however, much more than a concern for recycling as opposed to disposal. The circular economy examines what can be done right along the supply-and-use chain so that as few resources as possible are used, then (and this is the important bit) to recover and regenerate products at the end of their conventional life. This means designing products for longevity, reparability, ease of dismantling and recycling.

Typical of the companies that have adopted this idea is Newlife Paints, on the south coast of England; it 'remanufactures' waste water-based paint back into a premium-grade emulsion. All products in the company's paint range guarantee a minimum of 50 per cent recycled content, made up from waste paint diverted from landfill or incineration. Industrial chemist Keith Harrison started the company. His wife encouraged him to clean up his unruly garage after many years of do-it-yourself projects. After realising that the stacked-up tins of paint represented a shocking waste, he began to search for a sensible and environmentally responsible solution to waste paint. '*I kept thinking I could do something with it, the paint had an intrinsic value. It would have been a huge waste just to throw it away.*' After two years of research, Keith successfully developed his technology, which involves removing leftover paint from tins that have been diverted from landfill, and blending and filtering them to produce colour-matched new paints. The company has also launched a premium brand, aimed at affluent customers with a green conscience, called Reborn Paints – a development that was partly funded by Akzo Nobel, maker of Dulux Paints. Keith now licenses his technology to companies, including the giant waste company Veolia. '*By licensing we can have more impact and spread internationally*,' he said. He also points out that manufacturers could plan more imaginatively for the afterlife of their products. For example, simply adding more symbols to packs to assist sorting waste paints into types would help.

sustainable sources of energy. Some innovations focus on making products easier to recycle or remanufacture once they have reached the end of their life. Some food packaging is designed to break down easily when disposed of, allowing its conversion into high-quality compost.

3.3 Diagnostic question: *Is the product and service innovation process defined?*

The final design of any service or product is the result of a journey from an initial idea through to the final offering. On this journey, a design will pass through a number of (usually defined) stages. Even products and services that are normally considered purely 'artistic', with a large element of creativity, such as movies or theatrical productions, are actually progressed through clearly defined stages. For example, the process of developing video games follows three broad phases: pre-production, production and post-production, each of which also follows a sequence of steps, as shown in Figure 3.8.

Like any innovation and design process, the video game design process will involve some blurring of the boundaries between stages, and often significant recycling and rework (see the case study at the end of this chapter). Furthermore, all such processes tend to move from a vaguely defined idea that is then refined and made progressively more detailed until it contains sufficient information for turning into an actual service, product or process. At each stage in this

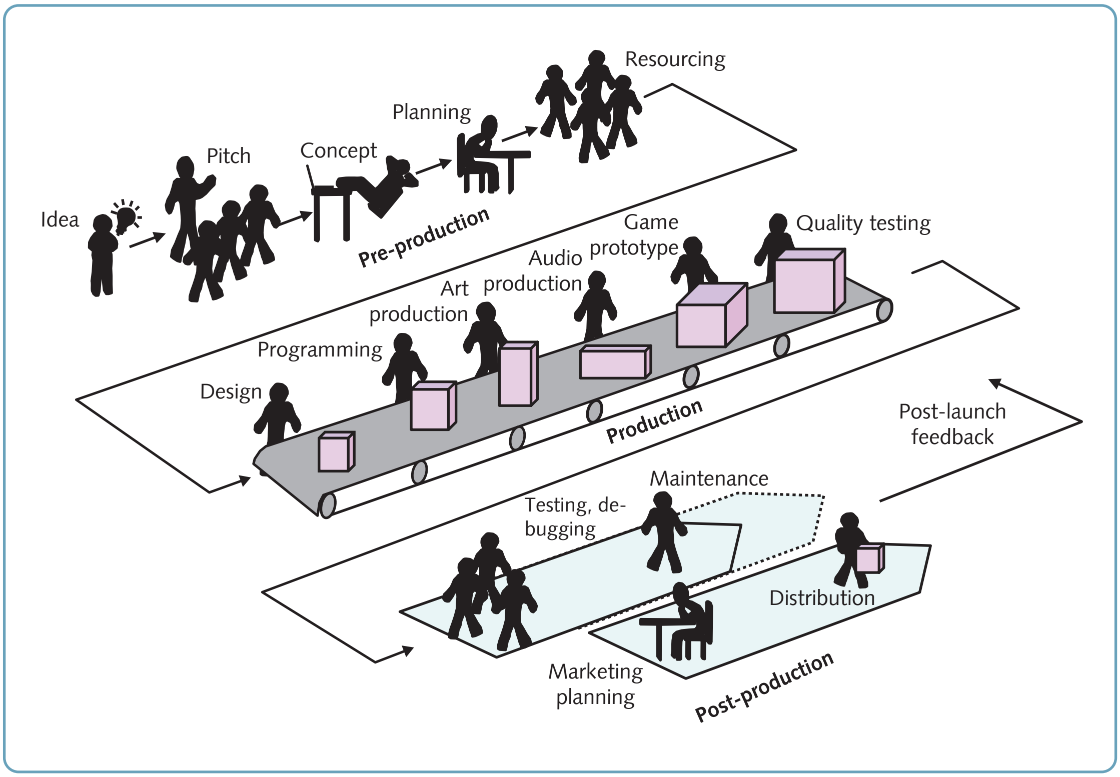


Figure 3.8 All products and services are created by defined operations processes, even highly creative processes such as video game development

> OPERATIONS PRINCIPLE
> ***Innovation processes involve a number of stages that move an innovation from a concept to a fully specified state.***

progression, the level of certainty regarding the final design increases as design options are discarded. The final design will not be evident until the very end of the process. Yet, many of the decisions that affect the eventual cost of delivery are made relatively early. For example, choosing to make a smartphone's case out of a magnesium alloy will be a relatively early decision that may take little investigation. Yet, although accounting for a small part of the total design budget, this decision may go a long way to determining the final cost of the phone. The difference between the 'budget spend' of the innovation process and the actual costs committed by the innovation process are shown in Figure 3.9, which also indicates the generic stages of product or service innovation. Although not all companies use these exact stages and there are often rework loops involved, there is still considerable similarity between the stages and sequence of the innovation process.

Concept generation

Concept generation is all about ideas, and ideas can come from anywhere. Often the expectations within organisations are that ideas will emerge from the research and development (R&D) or market research departments. However, this ignores the huge potential of other internal sources of innovation. Front-line service providers, in particular, are able to provide deep insights into what customers require based on informal interactions. Similarly, while many customer complaints are dealt with at a relatively operational level, they have the potential to act as a useful source of customer opinion within the innovation process. Suppliers can also be valuable in the innovation process because of their potential to improve the quality of products and services, minimise time to market, and spread the cost and risks of innovation.

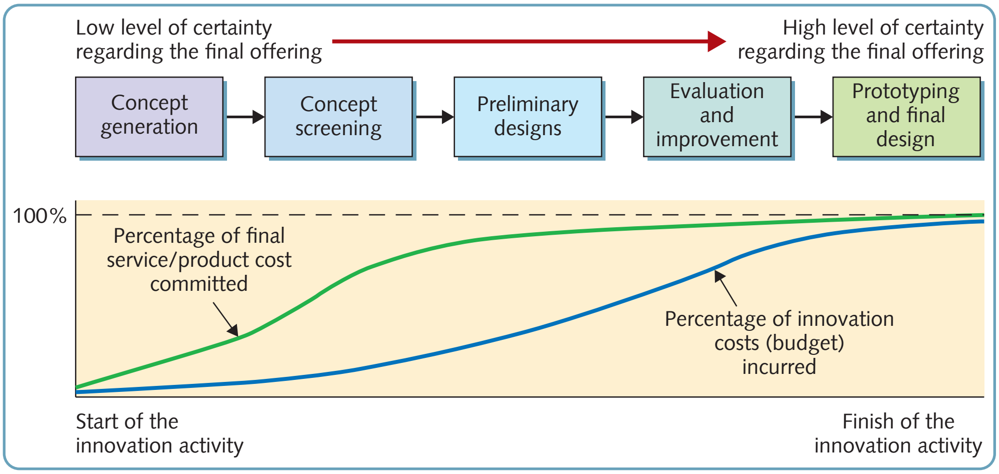


Figure 3.9 The stages in a typical innovation process and the design funnel effect – progressively reducing the number of possibilities until the final design is reached

'Lead users' and 'harbingers of failure'

A particularly useful source of customer-inspired innovation, especially for products and services subject to rapid change, are so-called 'lead users'. 'Lead users' are users who are ahead of the majority of the market on a major market trend, and who also have a high incentive to innovate. Producers seeking user innovations to manufacture try to source innovations from lead users, because these will be most profitable to manufacture. These customers, unlike most customers, have the real-world experience needed to problem solve and provide accurate data to inquiring market researchers. Since these lead users will be familiar with both the positives and negatives of the early versions of products and services, they are a particularly valuable source of potential innovative ideas.

If lead users can bring insights because of their expertise, by contrast another category of customer may be valuable because of their ability consistently to make bad purchase decisions. These customers have been termed 'harbingers of failure'. One study[5] claims that the same group of consumers has a tendency to purchase all kinds of failed products, time after time, flop after flop. As one of the authors of the study put it, 'These harbingers of failure have the unusual property that they keep on buying products that are taken from the shelves. These star-crossed consumers can sniff out flop-worthy products of all kinds. If you're the kind of person who bought something that really didn't resonate with the market, say, coffee-flavoured Coca-Cola, then that also means you're more likely to buy a type of toothpaste or laundry detergent that fails to resonate with the market.'

Ideas management

Obtaining new product or service ideas (or indeed any innovative ideas) from employees can also be a rich source of innovation. For example, the 3M Corporation has been highly successful in generating innovations by introducing formal incentives to encourage employee engagement. Employee-sourced ideas were traditionally done through paper-based 'suggestion schemes' where employees placed their ideas in a 'suggestion box'. Such schemes were often only partly effective, yielding few, low-quality ideas. Unless the running of the scheme was well resourced, it could be difficult to guarantee that all ideas were evaluated consistently and quickly. The scheme could lose credibility unless employees could track their ideas to confirm that they 'didn't just disappear'. However, the advent of 'ideas management' software tools has overcome some of these difficulties. Ideas management systems are a type of enterprise software (often web-based) that can help operations to collect ideas from employees, assess them and, if appropriate, implement them quickly and efficiently. Such systems can track ideas all the way through from inception to implementation, making it much easier to understand important performance measures such as where ideas are being generated, how many ideas submitted are actually implemented, the estimated cost savings from submitted ideas and any new revenues generated by implemented ideas. Often ideas management systems focus ideas on specific organisational targets and objectives, which it is claimed improves both the quality and quantity of ideas when compared with 'open' suggestion schemes.

Concept screening

Concept screening is the first stage of implementation where potential innovations are considered for further development. It is not possible to translate all concepts into viable product-service packages. Organisations need to be selective. For example, DuPont estimates that the ratio of concepts to marketable offerings is around 250:1. In the pharmaceuticals industry, the ratio is closer to 10,000:1. The purpose of concept screening is to take initial concepts and evaluate them for their feasibility (can we do it?), acceptability (do we want to do it?) and vulnerability (what are the risks of doing it?). Concepts may have to pass through many different screens, and several functions might be involved. Table 3.1 gives typical feasibility, acceptability and vulnerability questions for marketing, operations and finance functions.

During concept screening a key issue to consider is deciding how big the innovation should be and where it should focus – innovation to the customer offering as opposed to innovation to the process of delivery. The vast majority of innovation is continuous or incremental in nature. Here the emphasis is on steady improvement to existing offerings and to the processes that deliver them. This kind of approach to innovation is very much reflected in the lean and total quality management perspectives. On the other hand, some innovation is discontinuous and involves radical change that is 'new to the world'. Discontinuous innovation is relatively rare – perhaps 5–10 per cent of all innovations could be classified as such – but creates major challenges for existing players within a market. This is because organisations are often unwilling to disrupt current modes of working in the face of a barely emerging market, but by the time the threat has emerged more fully it may be too late to respond. Clayton Christensen refers to this problem as the 'Innovator's Dilemma', which supports renowned economist Joseph Schumpeter's idea that innovation should be a process of 'creative destruction'.[6]

Preliminary design

Having generated one or more appropriate concepts, the next stage is to create preliminary designs. For service-dominant offerings, this may involve documentation in the form of job instructions or 'service blueprints'. For product-dominant offerings, preliminary design involves

Table 3.1 Some typical evaluation questions for marketing, operations and finance

Evaluation criteria	*Marketing*	*Operations*	*Finance*
Feasibility	Is the market likely to be big enough?	Do we have the capabilities to deliver it?	Do we have access to finance to develop and launch it?
Acceptability	How much market share could it gain?	How much will we have to reorganise our activities to deliver it?	How much financial return will there be on our investment?
Vulnerability	What is the risk of it failing in the market-place?	What is the risk of us being unable to deliver it acceptably?	How much money could we lose if things do not go to plan?

OPERATIONS PRINCIPLE
A key innovation objective should be the simplification of the design through standardisation, commonality, modularisation and mass customisation.

defining product specifications (McDonald's has over 50 specifications for the potatoes used for its fries) and the bill of materials, which details all the components needed for a single product. At this stage, there are significant opportunities to reduce cost through design simplification. The best innovations are often the simplest. Designers can adopt a number of approaches to reduce design complexity. These include standardisation, commonality, modularisation and mass customisation.

Standardisation

Standardisation in product or service design is usually an attempt to overcome the cost of high variety by formally restricting it. Examples include fast-food restaurants specifying exactly how a prepared dish should look by providing a picture to its chefs; call centres providing on-screen scripts for operators to answer callers' questions; and although everybody's body shape is different, garment manufacturers produce clothes in a limited number of sizes (the range of sizes is chosen to give a reasonable fit for most, but not all, body shapes). Controlling variety is an important issue for most businesses, and is often built into the design of a product or service. For example, many aspects of service in global hotel chains are specified in standard operating procedures (SOPs). The Marriott hotel chain says that the company's prosperity rests on such things as its '66 Steps to Clean a Room' manual.[7]

Commonality

Common elements are used to simplify design complexity. If different services and products can draw on common components, the easier it is to deliver them. An example of this is Airbus, the European aircraft maker, which designed its aircraft with a high degree of commonality using fly-by-wire technology. This meant that ten aircraft models featured virtually identical flight decks, common systems and similar handling characteristics. The advantages of commonality for the airline operators include a much shorter training time for pilots and engineers when they move from one aircraft to another. This offers pilots the possibility of flying a wide range of routes from short-haul to ultra-long-haul and leads to greater efficiencies because common maintenance procedures can be designed with maintenance teams capable of servicing any aircraft in the same family. In addition, when up to 90 per cent of all parts are common within a range of aircraft, there is a reduced need to carry a wide range of spare parts. Similarly, Hewlett-Packard and Black & Decker use common platforms to reduce innovation costs.

Modularisation

This is a method of balancing two opposite forces: standardisation and customisation. It involves designing standardised 'sub-components' of an offering that can be put together in different ways. For example, the package holiday industry can assemble holidays to meet a specific customer requirement, from pre-designed and pre-purchased air travel, accommodation, insurance, and so on. Similarly, in education modular courses are increasingly used that allow 'customers' choice but permit each module to have economical volumes of students. Dell, a pioneer in computer manufacture, used the same logic for products, drawing together interchangeable sub-assemblies, manufactured in high volumes (and therefore lower cost), in a wide variety of combinations. In a similar way software engineering often involves modularisation to bring some degree of order to the development of large and complex pieces of software, which can often involve a large number of programmers. Here modularisation allows the sometimes thousands of lines of code to be broken up and organised by the tasks they perform.

Mass customisation

Flexibility in design can allow the ability to offer different things to different customers. Normally, high variety means high cost, but some companies have developed their flexibility in such a way that customised offerings are produced using high-volume processes and thus costs are minimised. This approach is called mass customisation. For example, Paris Miki, an upmarket eyewear retailer that has the largest number of eyewear stores in the world, uses its own 'Mikissimes Design System' to capture a digital image of the customer and analyse facial characteristics. Together with a list of each customer's personal preferences, the system then recommends a particular design and displays it on the image of the customer's face. In consultation with the optician, the customer can adjust shapes and sizes until they have chosen the final design. The frames are assembled within the store from a range of pre-manufactured components and the lenses ground and fitted to the frames. The whole process takes around an hour.

Evaluation and improvement

The purpose of this stage in the innovation process is to take a preliminary design and see if it can be improved before the offering is tested in the market. A number of techniques can be employed at this stage to evaluate and improve the preliminary design. Perhaps the best known is quality function deployment (QFD). The key purpose of QFD is to try to ensure that the eventual innovation actually meets the needs of its customers. It is a technique that was developed in Japan at Mitsubishi's Kobe shipyard and used extensively by Toyota, the motor vehicle manufacturer, and its suppliers. QFD is also known as the 'house of quality' (because of its shape) and the 'voice of the customer' (because of its purpose). The technique tries to capture what the customer needs and how it might be achieved. Figure 3.10 shows a simple QFD matrix used in the design of a promotional USB data storage pen. It is a formal articulation of how designers see the relationship between the requirements of the customer and the design characteristics of the offering.

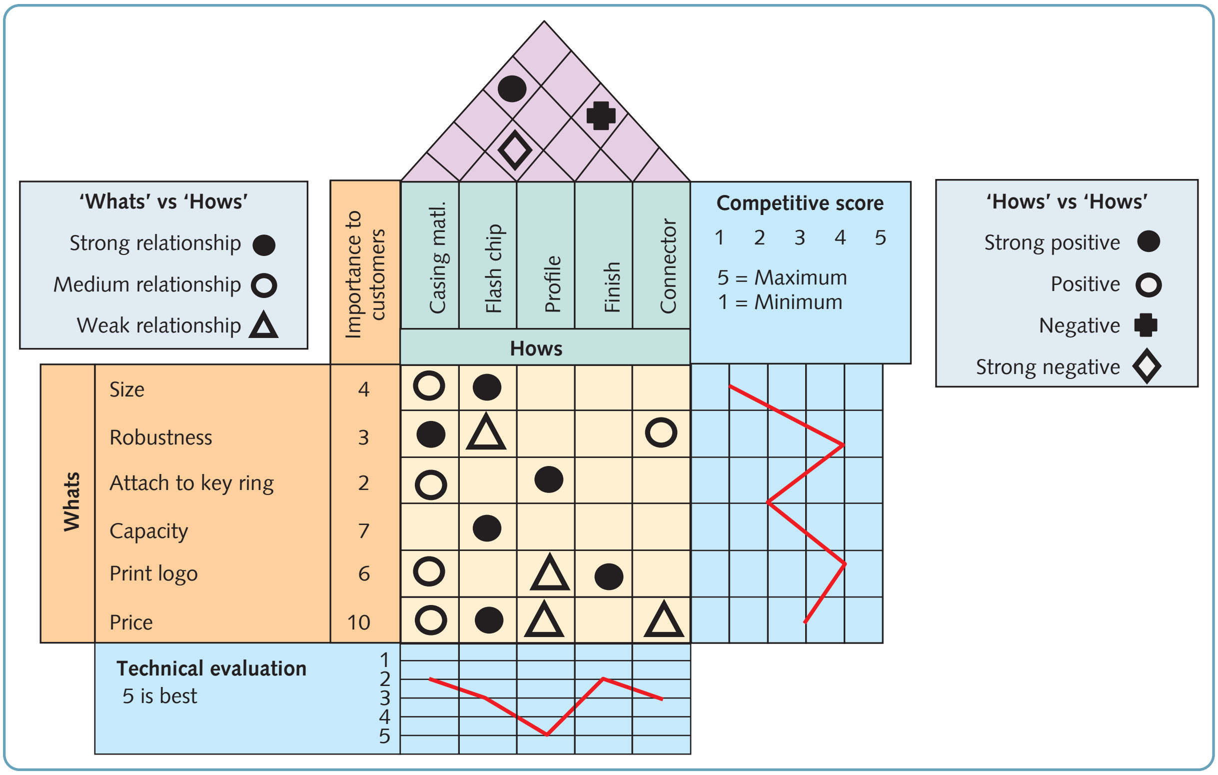


Figure 3.10 A simple QFD matrix for a promotional USB data storage pen

At this stage in the process, both creativity and persistence are needed to move from a potentially good idea to a workable design. One product has commemorated the persistence of its design engineers in its company name. In 1953 the Rocket Chemical Company set out to create a rust-prevention solvent and degreaser to be used in the aerospace industry. Working in their lab in San Diego, California, they made 40 attempts to work out the water displacing formula, so they called the product 'WD-40®', which literally stands for Water Displacement, 40th attempt. Originally used to protect the outer skin of the Atlas Missile from rust and corrosion, the product worked so well that employees kept taking cans home to use for domestic purposes. Soon after, the product was launched with great success into the consumer market.

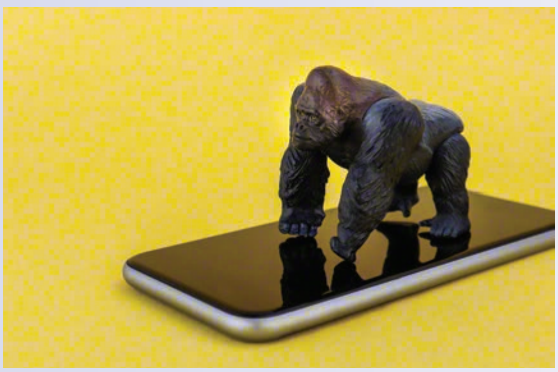
In fact, it's not just persistence that is important in the innovation process – failure itself may be beneficial if organisations can spot potential. Sometimes, when a design fails, it represents an opportunity to rethink the concept itself. For example, Pritt Stick, the world's first glue stick, was originally intended to be a super-glue, but product testing proved unsatisfactory. Henkel changed the product concept and successfully marketed the product as 'the non-sticky sticky stuff'. Similarly, a group of chemists working for the pharmaceutical giant Pfizer developed a new drug called 'Sildenafil'. Originally intended to help individuals with hypertension (high blood pressure) and angina, clinical trials of the drug proved unsuccessful, though doctors noticed a side effect of penile erection. Seeing the potential of this 'failed' innovation process, Pfizer marketed the drug as Viagra, for erectile dysfunction. In just two years, sales of Viagra had topped $1 billion and the product dominated the market.

Case example

Gorilla Glass[8]

Even if a new product or service idea does not make it to the end of a development process, the learning that's derived from the development process can still prove useful. This is a lesson well understood by Corning, one of the world's biggest glassmakers. For more than 165 years, it has combined its innovative ideas in glass science (and ceramics science and optical physics) with well-established manufacturing and engineering capabilities to develop its pioneering products and services. The scientists at Corning's research centre create thousands of new glass formulations and production innovations every year. Not all have immediate market potential. Those that may have potential are sent to the research centre's small glassworks for trial production, but even these may not make it to market. But Corning knows that the results of its research, even if it does not result in a new product, still has value. Everything learned from all its development projects is added to its 'knowledge bank'. Who knows, it could prove useful in the future.

This is exactly what happened during the development of Apple's revolutionary iPhone. Steve Jobs, the then boss (and founder) of Apple, had a request. Could Corning develop a faultlessly clear, strong and scratch-resistant glass that could be used to cover the screen of the newly designed iPhone, and please could they do it in six months? Staff at Corning's research centre delved into their files and found the results of a project from back in the 1960s. It recounted a project to develop a toughened lightweight glass for industrial use. Small volumes of the tough glass had been made, but the product was abandoned when few customers were interested. Corning reworked the formula to create a glass that was able to be produced thin and strong enough to be suitable for touch screens. (If the screen on a mobile phone is too thick, finger movement on the screen is harder to detect.) It called the new glass Gorilla Glass. As screens became thinner and mobile phones became skinnier, it went on to be improved progressively to retain its strength, and it has also found markets in other touch-screen applications.



WONGTANG/Shutterstock

Prototyping and final design

Product prototypes include everything from clay models to computer simulations. For more service-dominant offerings, prototyping may involve the actual implementation of the service on a pilot basis. For example, a retailer may organise the piloting of new services packages in a small number of stores in order to test customers' reaction to them. It may also be possible to 'virtually' prototype. For example, when designing a mass-transport hub, the behaviour and flow patterns of customers can be modelled using sophisticated simulation models, with impressive visuals that help to understand the effects of different layouts and volumes. In such simulations, safety is always an important objective, as would be efficient movement through the area. And although the means of prototyping may vary, the principle is always the same: do whatever one can to test out the innovation prior to delivery.

Alpha and beta testing

A distinction that originated in the software development industry but has spread into other areas is that between the alpha and beta testing of a product or service. Most software products include both alpha and beta test phases, both of which are intended to uncover 'bugs' (errors) in the product. Not surprisingly, alpha testing comes before beta testing. Alpha testing is essentially an internal process where the developers or manufacturers (or sometimes an outside agency they have commissioned) examine the product for errors. Generally, it is also a private process, not open to the market or potential customers. Although it is intended to look for errors that otherwise would emerge when the product is in use, it is, in effect, performed in a virtual or simulated environment, rather than in 'the real world'. After alpha testing, the product is released for beta testing; that is, when the product is released for testing by selected customers. It is an external 'pilot-test' that takes place in the real world (or near-real world, because it is still a relatively small, and short, sample) before commercial production. By the time a product gets to the beta stage most of the worst defects should have been removed, but the product may still have some minor problems that may only become evident with user participation. This is why beta testing is almost always performed at the user's premises without any of the development team present. Beta testing is sometimes called field testing, pre-release testing, customer validation, customer acceptance testing or user acceptance testing.

3.4 Diagnostic question: *Are the resources for developing product and service innovation adequate?*

> **OPERATIONS PRINCIPLE**
> ***For innovation processes to be effective they must be adequately resourced.***

For any process to operate effectively, it must be appropriately designed and resourced. Innovation processes are no different, so there are some strategic resourcing issues to consider – how much capacity to devote to innovation, how much of the innovation activity to outsource and what kinds of technology to use in the innovation process.

Is there sufficient innovation capacity?

Capacity management involves deciding on the appropriate level of capacity needed by a process and how it can be adjusted to respond to changes in demand. For innovation, demand is the number of new designs needed by the business. The chief difficulty is that, even in very large companies, the rate of innovation is not constant. This means that product and service design processes are subjected to uneven internal 'demand' for designs, possibly with several new offerings being introduced to the market close together while at other times little innovation is needed. This poses a resourcing problem because the capacity of an innovation activity is often difficult to flex. The expertise necessary for innovation is embedded within designers, technologists, market analysts and so on. It may be possible to hire-in some expertise as and when it is needed but much design resource is, in effect, fixed.

Such a combination of varying demand and relatively fixed design capacity means some organisations are reluctant to invest in innovation processes because they see it as an underutilised resource. A vicious cycle can develop, in which companies fail to invest in innovation resources simply because many skilled design staff cannot be hired in the short term, resulting in innovation projects overrunning or failing to deliver appropriate solutions. This, in turn, may lead to the company losing business or otherwise suffering in the market-place, which makes the company even less willing to invest in innovation resources.

Should all innovation activities be done in-house?

Just as there are supply networks that produce services and products, there are also supply networks of knowledge that connect suppliers and customers in the innovation process; this is sometimes called the 'design (or development) network'. Innovation processes can adopt any position on a continuum of varying degrees of design engagement with suppliers, from retaining all the innovation capabilities in-house, to outsourcing all the innovation work. Between these extremes are varying degrees of internal and external capability. Figure 3.11 shows some of the more important factors that will vary depending on where an innovation process is on the continuum. Resources will be easy to control if they are kept in-house because they are closely aligned with the company's normal organisational structures, but control should be relatively loose because of the extra trust present in working with familiar colleagues. Outsourced design may give the impression of greater control, because of the use of penalty clauses for delay. Costs may also be seen in different ways depending on whether design is outsourced. External innovation tends to be regarded as a variable cost. In-house innovation is more of a fixed cost. One inhibitor to open innovation is the fear of knowledge leakage. Firms become concerned that experience gained through collaboration with a supplier of design expertise may be transferred to competitors. There is a paradox here. Businesses usually outsource design primarily because of the supplier's capabilities that are themselves an accumulation of specialist knowledge from working with a variety of customers. Without such knowledge 'leakage', the benefits of the supplier's accumulated innovation capabilities would not even exist.

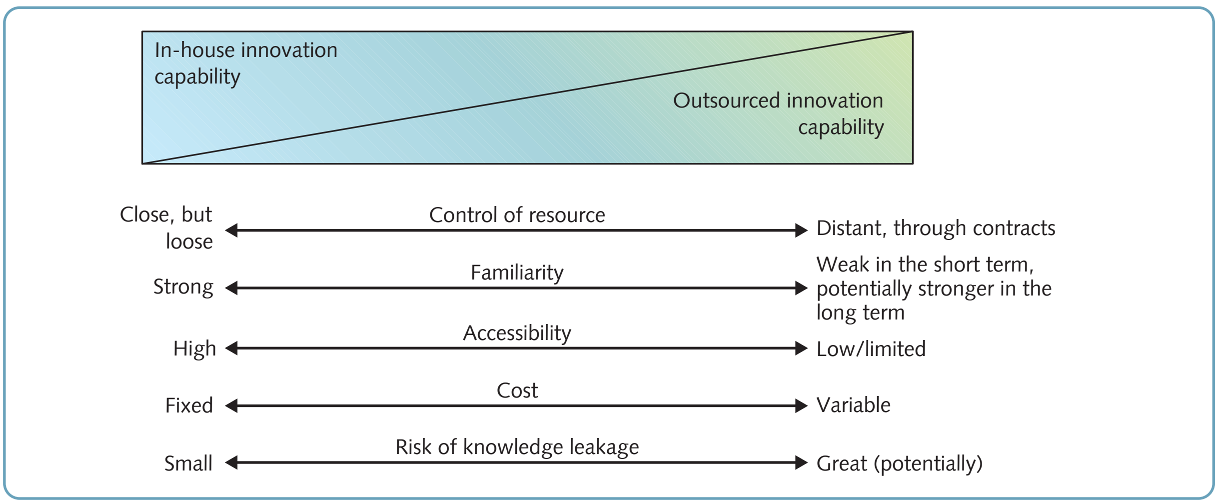


Figure 3.11 Some implications of the in-house–outsourced continuum

Open-source innovation – using a development community

The implication of the model of innovation that we presented in Figure 3.9 is that innovation is largely an internal affair, with ideas and their subsequent development into marketable products and services happening exclusively inside the organisation. However, in recent years there is evidence that some firms are more likely to widen their scope to include resources from outside in the innovation process. This may be formalised and could include a wide community of individuals, suppliers, customers, research institutes (such as universities) and even commercial rivals. This is generally referred to as 'open innovation'.[9] And 'openness' can work in two ways. What is sometimes called 'inbound' open innovation is where an organisation deliberately promotes external input from collaborators. 'Outbound' open innovation is where organisations permit unexploited ideas to go outside for others to use.

Arguably, the best-known examples of open innovation have come from the software industry. An open community, including the people who use the products, develops many of the software applications that we all use. If you use Google, Wikipedia or Amazon, you are using open-source software. The basic concept of open-source software is simple: large communities of people around the world, who have the ability to write software code, come together and produce a software product. The finished product is not only available to be used by anyone or any organisation for free but is regularly updated to ensure it keeps pace with the necessary improvements. The production of open-source software is very well organised and, like its commercial equivalent, is continuously supported and maintained. However, unlike its commercial equivalent, it is absolutely free to use. Over the last decade, the growth of open source has been significant with many organisations transitioning over to using this stable, robust and secure software. Open source has been the biggest change in software development for decades and is setting new open standards in the way software is used. The open nature of this type of development also encourages compatibility between products.

More generally, open innovation has several other benefits:

- The exposure of an innovation can allow customers and competitors to explore alternative markets or applications for innovations.
- Products and services can be refreshed if partners are able to exploit underlying processes and technology.
- It helps establish an innovation community that can support customer engagement.

Crowdsourcing

Closely related to the open sourcing idea is that of 'crowdsourcing'. Crowdsourcing is the process of getting work, or funding, or ideas (usually online) from a crowd of people. Although in essence it is not a totally new idea, it has become a valuable source of ideas largely through the use of the internet and social networking. For example, Procter & Gamble, the consumer products company, asked amateur scientists to explore ideas for a detergent dye whose colour changed when enough had been added to dishwater. Similarly, the LEGO Ideas platform allows users to submit their ideas for new LEGO sets. They can also vote and offer feedback for ideas that have been submitted. If an idea gets over 10,000 votes it is reviewed by LEGO, who then work with the idea's originator. Other uses of crowdsourcing involve government agencies asking citizens to prioritise spending (or spending-cutting) projects.

Is appropriate technology being used in the innovation process?

Technology has become increasingly important in innovation activities. Simulation software, for example, is now common in the design of everything from transportation services through to chemical factories. It allows developers to make design decisions in advance of the creation of

Case example

BT's open innovation ecosystem[10]

BT is one of the world's leading providers of communications services and solutions, with customers in 180 countries. Its primary activities include the provision of networked IT services globally; local, national and international telecommunications services, TV and internet products and services; and converged fixed-mobile products and services. This is a business that thrives on innovation, and indeed BT has a world-renowned research and development organisation. But it is the way that the group has grown a broad ecosystem of innovation, spanning both internal and external partners, that has allowed BT to innovate in a way that has made it a world leader in developing and exploiting innovation. Innovation starts with ideas, and ideas can come from many different sources. BT's 'open innovation' model is open in the broadest sense. It means working and collaborating with actual and potential suppliers, customers and its own staff. In fact, many of the novel ideas and technologies on which BT depend are created both within its own formal research and development facilities and by its staff more widely. The group's 'new ideas scheme' has an important role in the company's open innovation drive. It acknowledges and exploits the creativity of the group's 90,000-plus employees worldwide. They can use the scheme to propose ideas that could help with innovation in BT's processes, or improve its products and services. Staff can submit informal 'Eureka moment' ideas, or more carefully developed concepts. The ideas are then reviewed by experts in the business, and those ideas with the biggest potential are taken further along the development process. Of the approximately 2,000 ideas submitted every year around 50 make it through to implementation.

But it is the way that BT is 'open' to external innovation that has attracted much attention. In the world of telecoms, innovation is distributed around a number of innovation 'hot-spots' including Silicon Valley, Israel, Japan, Korea, Singapore, Hong Kong, India, China and throughout Europe. The central thrust of BT's global open innovation model comes from its technology scouting unit. This has been advancing the group's innovation efforts since 2000, when its then Senior VP of Technology and Innovation was sent to Palo Alto to gain insights into emerging technologies and business models. Significantly, the head of BT External Innovation is still located not in the UK, but in Silicon Valley. At the heart of its network, the BT Infinity Lab in London allows it to keep track and often co-innovate with start-ups while innovation scouting teams scan the world of technology start-ups, venture capitalists and researchers for ideas. It is constantly looking for and evaluating ideas – everything from novel technologies, developments in how markets are developing, inventive operations processes, to pioneering business models. BT believes that its dedicated scouting capability amplifies the innovation-generating abilities by several orders of magnitude. BT's open innovation model also includes long-standing research partnerships with leading universities and business schools around the world including collaborations with the Massachusetts Institute of Technology (MIT), Cambridge University, Tsinghua University in Beijing and the group's joint research centre in the UAE (EBTIC).

the actual product or service. Designers can work through the experience of using the service or product and learn more about how it might operate in practice. They can explore possibilities, gain insights and, most importantly, they can explore the consequences of their decisions. Innovation technologies are particularly useful when the design task is highly complex, because they allow developers to reduce their own uncertainty of how services or products will work in practice. Technologies also consolidate information on what is happening in the innovation process, thus presenting a more comprehensive vision within the organisation.

Generative design

Like almost all aspects of operations and process management, product and service innovation is not immune from the increased use of technology. One particularly useful technology is the use of 'generative design' as an approach to exploring alternative designs. It involves designers specifying important design goals, parameters and performance requirements, after which generative design software explores all the possible permutations of a solution, from which it generates design alternatives. More advanced versions use artificial intelligence to learn from each iteration. Its proponents describe generative design as exploiting machine learning to mimic nature's evolutionary approach to design. One of its obvious advantages is speed: digital simulation and analysis can evaluate potential designs in seconds. But it can also generate design options that would not otherwise have been thought of. Generative design software can explore a much larger range of possible solutions, comparing the results of thousands of simulations to close in on a design that delivers the best combination of attributes. A wide variety of companies are using it to solve design problems and derive solutions that human designers would find difficult to come up with on their own. These include Airbus (aerospace), Under Armour (sports garments and equipment) and Stanley Black & Decker (hand tools).[11]

Computer-aided design (CAD)

Probably the best-known (and well-established) innovation technology is computer-aided design (CAD). CAD systems store and categorise component information and allow designs to be built up on screen, often performing basic engineering calculations to test the appropriateness of proposed design solutions. They provide the ability to create a modified product drawing and allow for the swift addition of conventionally used shapes to the computer-based representation of a product. Designs created on screen can be saved and retrieved for later use, which enables a library of standardised parts and components to be built up. Not only can this dramatically increase the productivity of the innovation process, it also aids the standardisation of parts in the design activity.

Digital twins

Michael Grieves of Florida Institute of Technology's Centre for Lifecycle and Innovation Management popularised the term 'digital twin'.[12] It is the combination of data and intelligence that represent the structure, context and behaviour of a physical system of any type, offering an interface that allows one to understand past and present operation, and make predictions about the future. In other words, it is a powerful digital 'replica' that can be used instead of the physical reality of a product. Using the digital twin rather than the real product can significantly improve its operational performance without the expense of working on the real thing. Not only that, but the use of the digital twin can extend throughout the product's life to provide valuable information to its user and evidence on how it is actually performing. Thus, products live two parallel lives: one in real life, one in digital form. For example, digital twins can monitor and simulate possible future scenarios and predict the need for repairs and other problems before they occur. This allows design improvements to be made before a product is used by customers, as well as during its lifespan.

Knowledge management technologies

In many professional service firms, such as management consultancies, design involves the evaluation of concepts and frameworks that can be used in client organisations to diagnose problems, analyse performance and construct possible solutions. They may include ideas of industry best practice, benchmarks of performance within an industry, and ideas that can be transported across industry boundaries. However, management consulting firms are often geographically dispersed and staff are often away from the office, spending most of their time in client organisations. This creates a risk for such companies of 'reinventing the wheel' continually. Most consultancy companies attempt to tackle this risk by using knowledge management routines based on their intranet capabilities. This allows consultants to put their experience into a common pool, contact other staff within the company who have skills relevant to a current assignment and identify previous similar assignments. In this way, information is integrated into the ongoing knowledge innovation process within the company and can be tapped by those charged with developing innovations.

3.5 Diagnostic question: *Is the development of products and services and of the process that created them simultaneous?*

Every product and service innovation has to be created, so it is a mistake to separate the design of products and services from the design of the processes that will deliver them. Merging these two processes is sometimes called simultaneous (or interactive) design. Its key benefit is to reduce the elapsed time taken for the whole innovation activity. As noted earlier, reducing time-to-market (TTM) can give an important competitive advantage. If it takes a business two years to develop an offering with a given set of resources, it can only introduce new offerings every two years. If its rival can develop offerings in one year, it can introduce innovations into the market twice as often. This means the rival does not have to make such radical improvements in performance each time it introduces a new offering because it is introducing them more frequently. The factors that can significantly reduce time-to-market for innovations include the following:

> OPERATIONS PRINCIPLE
> ***Effective simultaneous innovation reduces time to market.***

- Integrating the design of the product–service offerings and the design of the process used to create and deliver them.
- Overlapping the stages in the innovation process.
- An early deployment of strategic decision-making and resolution of design conflict.
- An organisational structure that reflects the nature of the offering.

Integrating the design of the offering and design of the process

What looks good on screen may prove difficult to create and deliver in practice. Conversely, a process designed for one set of services or products may be incapable of creating different ones. It clearly makes sense to design offerings and operations processes together. For services, organisations have little choice but to do this because the process of delivery is usually part of the offering. However, it is also useful to integrate the design of the offering and the process for products, but there are real barriers to doing it. First, the timescales involved can be very different. Offerings may be modified, or even redesigned, relatively frequently. The processes used to create and deliver an offering may be far too expensive to modify every time the offering changes. Second, the people involved with the innovation on the one hand, and operations process design on the other, are likely to be organisationally separate. Finally, it is sometimes not possible to design an ongoing process for the creation and delivery of services and products until they are fully defined.

Yet none of these barriers is insurmountable. Although it may not be possible to change ongoing processes every time there is a change to the offering, they can be designed to cope with a range of potential services and products. The fact that design staff and operations staff are often organisationally separate can also be overcome. Even if it is not sensible to merge the two functions, there are communication and organisational mechanisms to encourage the two functions to work together. Even the claim that ongoing processes cannot be designed until the nature of the offering is known is not entirely true. There can be sufficient clues emerging from innovation activities for process design staff to consider how they might modify ongoing processes. This is a fundamental principle of simultaneous design, considered next.

Overlapping the stages of the innovation process

We have described the innovation process as a set of individual, predetermined stages, with one stage being completed before the next one commences. This step-by-step, or sequential, approach has been commonly applied in many organisations. It has some advantages. It is easy to manage and control innovation processes organised in this way because each stage is clearly defined. In addition, as each stage is completed before the next stage is begun, each stage can focus its skills and expertise on a limited set of tasks. The main problem of the sequential approach is that it is both time-consuming and costly. When each stage is separate, with a clearly defined set of tasks, any difficulties encountered during the design at one stage might necessitate the design being halted while responsibility moves back to the previous stage. This sequential approach is shown in Figure 3.12(a).

Often there is really little need to wait until the absolute finalisation of one stage before starting the next. For example, perhaps while generating the concept, the evaluation activity of screening and selection could be started. It is likely that some concepts could be judged as 'non-starters' relatively early on in the process of idea generation. Similarly, during the screening stage, it is likely that some aspects of the design will become obvious before the phase is finally complete. Therefore, the preliminary work on these parts of the design could commence at that point. This principle can be taken right through all stages, each stage commencing before the previous one has finished, so there is simultaneous or concurrent work on the stages, as depicted in Figure 3.12(b).

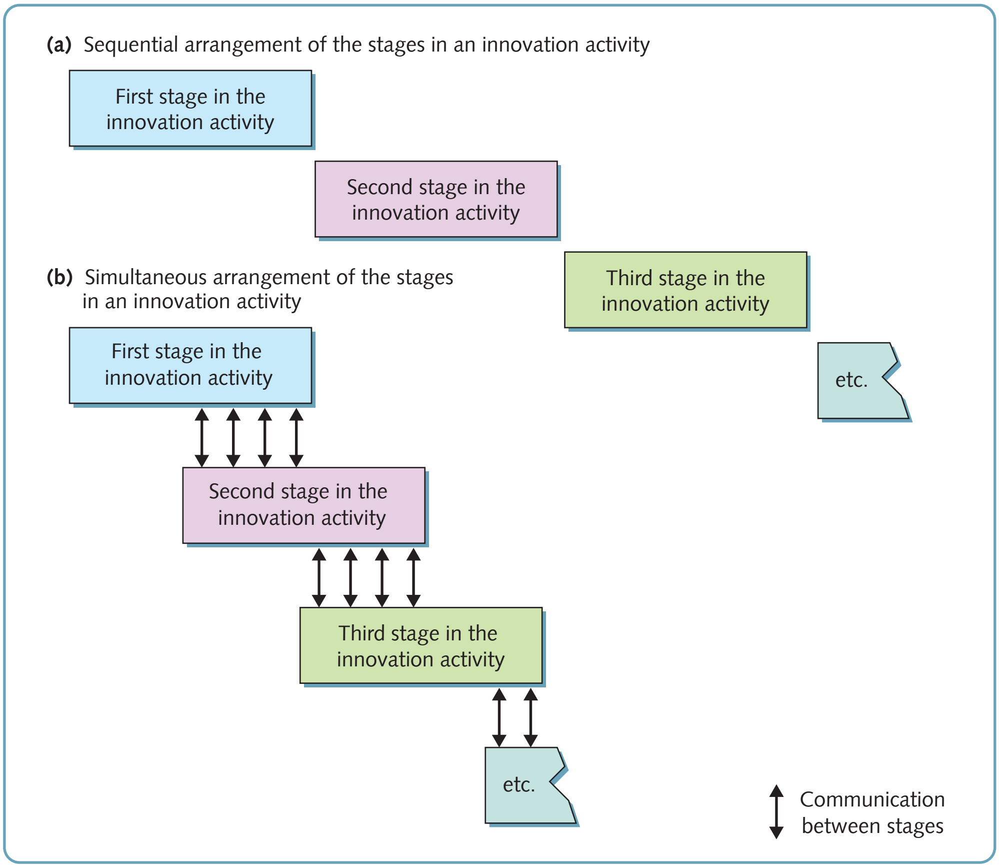


Figure 3.12 Sequential (a) and simultaneous (b) arrangements of the stages in an innovation activity

We can link this idea with the idea of uncertainty reduction, discussed earlier, when we made the point that uncertainty reduces as the design progresses. This also applies to each stage of innovation. If this is the case, there must be some degree of certainty that the next stage can take as its starting point, prior to the end of the previous stage. In other words, designers can continually react to a series of decisions and clues given to them by those working on the preceding stage. However, this can work only if there is effective communication between each pair of stages.

Deploying strategic intervention and resolving conflicts early

A design decision, once made, need not shape the final offering irrevocably. All decisions can be changed, but it becomes increasingly difficult to do so as the innovation process progresses. At the same time, early decisions are often the most difficult to make because of the high level of uncertainty surrounding what may or may not work as a final design. This is why the level of debate, and even disagreement, over the characteristics of an offering can be at its most heated in the early stages of the process. One approach is to delay decision-making in the hope that an obvious 'answer' will emerge. The problem with this is that if decisions to change are made later in the innovation process, these changes will be more disruptive than if they are made early on. The implication of this is that first, it is worth trying to reach consensus in the early stages of the innovation process even if this seems to be delaying the total process in the short term, and second, there is a particular need for strategic intervention into the innovation process by senior management during these early stages. Unfortunately, there is a tendency for senior managers, after setting the initial objectives of the innovation process, to 'leave the details' to technical experts. They may only become engaged with the process again in the later stages as problems start to emerge that need reconciliation or extra resources.

> **OPERATIONS PRINCIPLE**
> ***The innovation process requires strategic attention early, when there is most potential to affect design decisions.***

Organising innovation processes in a way that reflects the nature of the offering

The innovation process will almost certainly involve people from several different areas of the business that will have some part in making the decisions shaping the final offering. Yet any design project will also have an existence of its own. It will have a project name, an individual manager or group of staff who are championing the project, a budget and, hopefully, a clear strategic purpose within the organisation. The organisational question is which of these two ideas – the various organisational functions that contribute to the innovation or the project itself – should dominate the way in which the design activity is managed?

There is a range of possible organisational structures – from pure functional to pure project forms. In a purely functional organisation, all staff associated with the innovation project are based unambiguously in their functional groups. There is no project-based group at all. They may be working full-time on the project but all communication and liaison are carried out through their functional manager. At the other extreme, all the individual members of staff from each function involved in an innovation project could be moved out of their functions and perhaps even co-located in a task force dedicated solely to the project. A project manager, who might hold the entire budget allocated to the innovation project, could lead the task force. Not all members of the task force necessarily have to stay in the team throughout the design period, but a substantial core might see the project through from start to finish. Some members of a design team may even be from other companies. In-between these two extremes there are various types of matrix organisation with varying emphasis on these two aspects of the organisation (see Figure 3.13):

- *Functional organisation* – the innovation project is divided into segments and assigned to relevant functional areas and/or groups within functional areas. The project is co-coordinated by functional and senior management.

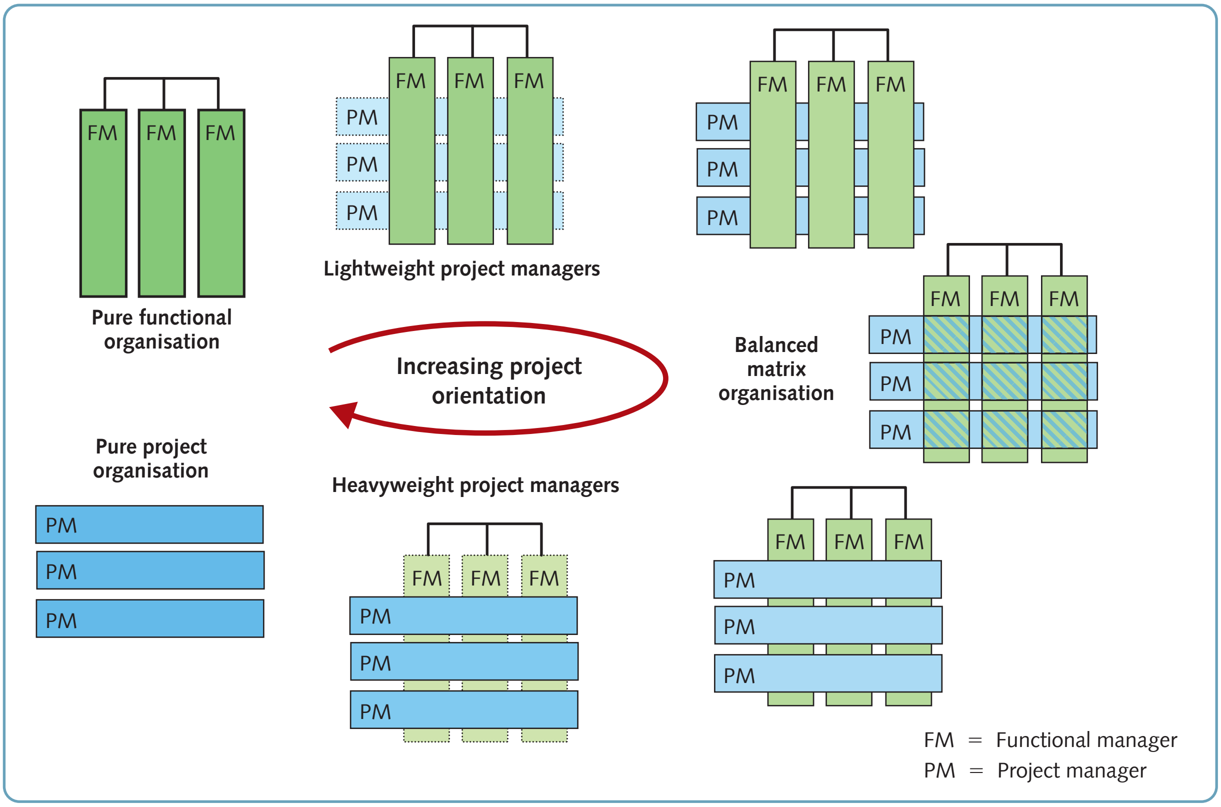


Figure 3.13 **Organisational structures for innovation processes**

- *Functional matrix (or lightweight project manager)* – a person is formally designated to oversee the project across different functional areas. This person may have limited authority over the functional staff involved and serves primarily to plan and coordinate the project. Functional managers retain primary responsibility for their specific segments of the project.
- *Balanced matrix* – a person is assigned to oversee the project and interacts on an equal basis with functional managers. This person and the functional managers work together to direct innovation activities and approve technical and operational decisions.
- *Project matrix (or heavyweight project manager)* – a manager is assigned to oversee the project and is responsible for its completion. Functional managers' involvement is limited to assigning personnel as needed and providing advisory expertise.
- *Project team (or tiger team)* – a manager is given responsibility for a project team composed of a core group of personnel from several functional areas assigned on a full-time basis. The functional managers have no formal involvement.

Although there is no clear 'winner' among the alternative organisational structures, there is increasing support for structures towards the project rather than functional end of the continuum. Some authorities argue that heavyweight project manager structures and dedicated project teams are the most efficient forms of organisation in driving competitiveness, shorter lead times and technical efficiency.

Perhaps of more interest is the suitability of the alternative structures for different types of innovation. Matrix structures are generally deemed to be appropriate for both simple and highly complex projects. Dedicated project teams, on the other hand, are seen as appropriate for projects with a high degree of uncertainty, where their flexibility becomes valuable. Functionally based design structures, with resources clustered around a functional specialism, help the development of technical knowledge. Some organisations do manage to capture the deep technological and skills development advantages of functional structures, while at the same time co-coordinating between the functions to ensure satisfactory delivery of new service and product ideas. Perhaps the best-known of these organisations is Toyota. It has a strong functionally based organisation to develop its offerings. It adopts highly formalised development procedures to communicate between functions and places strict limits on the use of cross-functional teams. What is really different is its approach to devising an organisational structure for innovation that is appropriate for the company. The argument that most companies have adopted to justify cross-functional project teams goes something like this: 'Problems with communication between traditional functions have been the main reasons for failing to deliver new innovation ideas to specification, on time and to budget. Therefore let us break down the walls between the functions and organise resources around the individual development projects. This will ensure good communication and a market-oriented culture.' Toyota, on the other hand, has taken a different approach. Its argument goes something like this: 'The problem with cross-functional teams is that they can dissipate the carefully nurtured knowledge that exists within specialist functions. The real problem is how to retain this knowledge on which our future innovation depends, while overcoming some of the traditional functional barriers that have inhibited communication between the functions. The solution is not to destroy the function but to devise the organisational mechanisms to ensure close control and integrative leadership which will make the functional organisation work.'

Critical commentary

- The whole process-based approach to innovation could be interpreted as implying that all new offerings are created in response to a clear and articulated customer need. While this is usually the case, especially for services and products that are similar to (but presumably better than) their predecessors, more radical innovations are often brought about by the innovation itself creating demand. Customers don't usually know that they need something radical. For example, in the late 1970s people were not asking for microprocessors; they did not even know what they were. An engineer in the USA improvised them for a Japanese customer who made calculators. Only later did they become the enabling technology for the PC and, after that, the innumerable devices that now dominate our lives. Similarly, fly-by-wire, digital cameras, Maersk's super-slow container ships, sushi on conveyor belts and the iPad are all examples of innovations that have been 'pushed' by firms rather than 'pulled' by pre-existing customer demand.
- Nor does everyone agree with the dominant rational model in which possible design options are progressively reduced stage by stage through the optimisation of known constraints and objectives. For some, this neat model of the innovation, which underlies much business and engineering design literature, fails to accurately reflect the creativity, arguments and chaos that sometimes characterise real innovation projects. First, they argue, managers do not start out with an infinite number of options. No one could process that amount of information – and anyway, designers often have some set solutions in mind, looking for an opportunity to be used. Second, the number of options being considered often increases as time goes by. This may actually be a good thing, especially if the activity was unimaginatively specified in the first place. Third, the real process of innovation involves cycling back, often many times, as potential solutions raise fresh questions or become dead ends, and as requirements and constraints evolve. In summary, the idea of the design funnel does not describe the process of innovation, nor does it necessarily even describe what should happen. The action-centric or co-evolution perspective of innovation represents the antithesis of the rational model. It posits that offerings are designed through a combination of emotion and creativity; that the process by which this is done is generally improvised; and that the sequencing of stages is not universal in innovation processes.

SUMMARY CHECKLIST

- ☐ Is the importance of innovation as a contributor to achieving strategic impact fully understood?
- ☐ Is innovation really treated as a process?
- ☐ Is the innovation process itself designed with the same attention to detail as any other process?
- ☐ Are innovation objectives specified to give a clear priority between quality, speed, dependability, flexibility, cost and sustainability?
- ☐ Are the stages in the innovation process clearly defined?
- ☐ Are ideas and concepts for new offerings captured from all appropriate internal and external sources?
- ☐ Are potential offerings screened in a systematic manner in terms of their feasibility, acceptability and vulnerability?
- ☐ During preliminary design, have all possibilities for design standardisation, commonality and modularisation of design elements been explored?
- ☐ Has the concept of mass customisation been linked to the innovation process?
- ☐ Are potential offerings thoroughly evaluated and tested during the innovation process?
- ☐ Are the resources for developing innovation adequate?
- ☐ Is sufficient capacity devoted to the innovation process?
- ☐ Have all options for outsourcing parts of the innovation process been explored?
- ☐ Has the possibility of involving customers formally in the innovation process been explored?
- ☐ Are appropriate technologies, such as CAD, digital twinning and knowledge management, being used in the innovation process?
- ☐ Are the design of the offering and the design of processes that create and deliver it considered together as one integrated process?
- ☐ Is overlapping (simultaneous) of the stages in the innovation process used?
- ☐ Is senior management effort deployed early enough to ensure early resolution of design conflict?
- ☐ Does the organisational structure of the innovation process reflect the nature of the offering?
- ☐ Are some functions of the business more committed to innovating new service and product offerings than others?
- ☐ If so, have the barriers to cross-functional commitment been identified and addressed?

Case study

Widescale Studios and the Fierybryde development

'*Anyone who has been involved with designing and constructing video games will tell you that game development never goes as planned. I sometimes think that it is a miracle that any game gets developed. Technical glitches, bottlenecks in production, conflicting creative egos, pressure from publishers, they will all throw you off course during the development cycle. It is a process that occupies the area on the borderline between art and technology. Yet, although video game development is an uncertain and complex process, it is how the development process is managed that is the key feature in whether a game will go on to be a success.*' (Izzy McNally, Co-Owner Widescale Studios)

Widescale Studios was a videogame development studio, located in the midlands of the UK. It had been founded seven years ago by Izzy McNally and Oli Chambers, when they left a larger studio to gain 'some creative independence'. Video game software development studios are the organisations that actually create the games. There are many thousands of such studios worldwide, some large, but most employing less than 30 people.[13] Some studios are owned by video game publishers, of which some also produce gaming hardware, and some, like Widescale, are independent. Publishers market and sell the games, manage relationships with distributors, platform providers and retailers, conduct market research and advertise games.

Originally from California, Izzy was, by background, an artist and writer. Oli started as a programmer, but had moved into becoming an executive producer. (In the industry, an executive producer is the person who is responsible for the overall coordination of the development.) Both Izzy and Oli admitted that their desire for more creative independence had not fully materialised: '*Since we started, we have been surviving as an independent studio by taking on contracts from the bigger studios, and we have built a good reputation. But if we don't have another contract ready to go when the last one finishes, we are in trouble. It can be dispiriting constantly looking for work to keep us afloat. That was why Fierybryde was so exciting.*' (Oli Chambers)

The Fierybryde project

Fierybryde was an idea for a role-playing game (RPG) that had come out of a number of brainstorming sessions between Izzy, Oli and Hussein Malik in the middle of an unusually warm and pleasant summer. Hussein was a developer and self-confessed 'fanatical gamer' who had joined Widescale soon after it was founded. A role-playing game is a video game in which players assume the roles of characters who are protagonists in a fictional setting. The senior team at Widescale were excited at the concept of Fierybryde, and saw it as an opportunity to develop a game of their own that would (potentially) give them both creative and financial independence. The Fierybryde concept was intriguing, although not totally novel. The game's setting was a combination of space exploration and 'wild west' adventure (Fierybryde was the name of the spaceship in the story) with various characters who possessed different skills and psychological traits. The purpose of the game was to build an intergalactic trading empire while avoiding interference from political and commercial rivals.

Traditionally, independent studios who wanted to develop a game such as Fierybryde had four methods of raising funds. First, they could pitch the idea to a publisher. Most video game development was funded by big publishers. However, publishers almost always insisted on terms that were more favourable to them than the developers. Second, the studio could seek private investors who would put their own money into the company and share any subsequent profits. The downside to this for Izzy and Oli would be a certain loss of independence. Third, the studio could attempt to raise money by crowdfunding, asking for (relatively small) donations from thousands of potential future users of the game in return for preferential access to the finished game. It was an increasingly popular method of raising funding, but limited to relatively small sums in total, often less than £1 million (the typical budget for an RPG would be tens of millions of pounds or more). Finally, the studio could start the development from their own saved capital, then fund the ongoing costs from the profits from their other work. This was the approach chosen by Widescale, who had a retained cash pot of around £700,000. If successful, Fierybryde could provide a stable stream of income, without substantial rights and royalties going to some big publisher. In turn, this would let the studio pursue more interesting projects in the future. Table 3.2 shows Widescale's projected cash flow forecast as of the start of the project.

The development process

Video game development is an uncertain and complex process, but a key feature in whether a game will go on to be a success is the way the development process is managed from concept through to launch. Although different studios use slightly different terms, game development is broken down into three stages: pre-production, production and post-production. Pre-production is the stage where the developers have to answer some fundamental questions about the game, including the market it is aimed at, the platform it will play on, the type of game it is going to be, the budget, the basic storyline and the timescale (at least nominally). The production phase is usually the most

Table 3.2 Widescale's planned total cash flow (£000)

Year	*1*		*2*				*3*				*4*	
Qtr	*3*	*4*	*1*	*2*	*3*	*4*	*1*	*2*	*3*	*4*	*1*	*2*
General revenue	2,500	2,500	2,300	2,500	2,000	2,400	2,000	2,000	2,000	2,000	2,000	2,000
General costs	1,600	1,600	1,600	1,600	1,600	1,600	1,600	1,600	1,600	1,600	1,600	1,600
Fierybryde revenue								1,500	4,500	3,500	1,500	1,500
Fierybryde costs	300	500	550	600	800	800	800	600	300	200	100	100
Cash	1,300	1,700	1,850	2,150	1,750	1,750	1,350	2,650	7,250	10,950	12,750	14,550

resource-intensive phase, and is often the phase that is the most uncertain and difficult to plan. It involves programmers, character artists, graphic designers, audio designers, voice actors, quality testers and producers who, in the words of Oli, '*provide the glue that makes it all happen*'. The aim of all of them is to make a game that will be original, fun and involving. They do this by using new gameplays, gripping underlying stories, enhanced graphics and convincing characters. If a game fails to meet users' concepts of 'quality play', they could readily switch to other games. Post-production manages the transition of the game into the market. Often a publisher will become involved at this stage, if they weren't before. Even at this stage, quality assurance continues because bugs in the software always continue to emerge. A 'hype' video extract from the game with a mix of graphics and sample gameplay will probably be released for marketing purposes, and a spot at one of the major gaming conventions may be arranged.

Starting the Fierybryde development

Work started on the outline of the Fierybryde concept over the summer, with Izzy drafting an outline script and Oli working on some technical issues such as the number of 'levels' the game should have and how many maps it should contain. The project was formalised with its own budget in September, when Hussein was asked to put together plans for how the development would progress. His first decision was to hire Ross Avery, who had been his boss at his previous studio. Ross had wide experience in the game development industry, largely in senior executive developer and producer roles.

Ross and Hussein formed the core of the Fierybryde team and were joined by planners and developers, both newly recruited and moved from Widescale's other work. However, Oli recognised that Fierybryde should not put any of the studio's regular projects at risk. '*They are our "bread and butter", each has a deadline and a budget that we must stick to. Fierybryde has more flexibility because it's directly under our control. Of course, we had a budget for it, but there was still considerable flexibility. So we looked at the budget and asked the question, do we want a team of 10 people working for 10 months or a team of 20 people working for five months? Theoretically, one could even have had a team of two people working for 50 months, but that would have been ridiculous. Also, at different times in the development process one will need different numbers of developers with specific skills. The balance was always between allocating the appropriate resources to Fierybryde without interfering with our other work.*'

By November it had become clear that Izzy would have to decide whether to take on responsibility for developing the games script herself or to hire-in a script writer. In the end she hired a part-time script writer who had experience of television work. Izzy admitted that it was a mistake: '*Script writing for a video game is totally different from writing for television or writing a novel. I underestimated this. In a script for television the narrative moves in a linear direction. With a video game the narrative is more like a tree. Each player can move along different branches of the tree depending on the decisions that they make. A script writer has to make up dialogue for many different scenarios, knowing that each individual player will see only one of them. It wasn't the fault of the writer we hired, it was my fault in underestimating the differences. In the end I had to take over far more of the script writing than I had intended.*'

The scripting and storyboarding of the game continued into the new year, but by January tensions had begun to emerge between Oli, who was concerned about the rate that the project was burning though the budget, and Ross, who wanted the script, characterisations and overall architecture of the game settled before the production phase commenced. Oli wanted to get the production stage of development started as soon as possible. '*Widescale's strength was in the actual production stage of development. That's what we spend most of our time doing. If we weren't good at keeping to schedule, we couldn't have survived as a contract developer. Also, I thought that we had an outline script and the overall structure of the game more or less sorted from mine and Izzy's work over the summer. I do understand that when a new person like Ross first joins the team, the temptation is to try to sit down at the beginning of the development process and settle the whole script from start to finish. But it has to develop naturally; developing the script for a game is essentially an iterative process*'.

The production phase

Although there were still uncertainties, and some disagreement around the game's storyline, by the end of January Oli had decided to formally move on to the production phase and allocated developers and artists to the project. He also started briefing the freelance graphics designers, sound designers and voice artists that they would need later in the process. In early February Ross resigned. He was philosophical about it: '*It's not unusual in this business. There will always be some tension between whoever is in charge of operationalising a concept and the studio owner. The important thing is who holds the budget? In this case the owners [Oli and Izzy] didn't want to give up control. It's their company and their money, so I guess they have the last word on any decisions, big or small. But, personally, I like to have more control than they wanted to give me.*'

From that point Oli acted as executive producer for the project, with Hussein overseeing technical issues and Izzy 'creative' ones. However, during the spring, the development fell increasingly behind schedule. Hussein admitted that many of the problems were the result of their own decisions. '*We started using a new 3D graphics package two months into the initial development. It allowed a new rendering approach that looked particularly exciting. It made the graphics better than we thought possible. It did give us some spectacular effects, but also gave us two problems. First, we totally underestimated the learning curve necessary to master the package. It took our developers a month or two to get used to the package and this delayed things more than we envisaged. Second, it became clear relatively quickly that the effect of the change was to knock the game's 'frame rate' down to the point where it looked poor. We knew the choice would affect frame rate but we just didn't anticipate the impact this would have on what the game felt like to play. Both these things undermined our ability to estimate how long some key stages of the development might take. Without the ability to estimate the individual development tasks it became particularly difficult to schedule the development as a whole*.'

By June, the development team were overcoming the problems with the new graphics package, when a further problem emerged. Hussein and Oli had decided to use a previously untried (by Widescale) game engine. (A game engine provides the software framework that allows developers to create video games.) Many commercial game engines are available to help game developers. Using one means that developers can focus solely on the logic of the game rather than getting bogged down in detail. A game engine allows code reuse, which usually means shorter development time and reduced cost. But not when a development team needs to learn it for the first time! Once again, progress slowed. Izzy believed that the problems were the result of trying to balance the eventual quality of the game with the costs of developing it. '*There is always something of a trade-off between the efficiency of the development process and the quality of the game that comes out of the process. Just to make things more complicated, you have to wait until the game is almost fully developed before you can judge quality, in terms of how much fun it is to play. So you have to manage the development process in the best way you think will promote both creativity and efficiency*.'

Both Izzy and Ollie liked the idea of giving as much freedom as possible to individual developers within the team. Compared to many studios the atmosphere was relaxed. On the whole it was thought that this had led to a good creativity that would eventually show through in the final game. Oli also thought that a more relaxed attitude helped to develop and retain the best development talent. '*Developers value flexibility to innovate and more ownership over the content they are working on, and some degree of independence from micromanagement. All too easily they can find another studio where they feel their skills are more valued.*'

However, there were times when it proved less than fully efficient. For example, at one time two developers each designed their own different versions of the same scene because their work had not been coordinated, costing several days of wasted effort. There were also problems in managing the studio's regular contract work alongside the Fierybryde project. At times the studio's other work took development resource away from the project. As one frustrated developer put it: '*It could be frustrating to suddenly have a colleague taken off the project for a week to work on another job, but hey, it was the revenue from this other work that was funding the project so naturally it took priority. Nevertheless, there were times when the other jobs were vacuuming up resources, and the whole development process was like being in a pressure cooker.*'

Project crunch and financial crunch

By November it was becoming obvious that Fierybryde was in trouble. It had fallen well behind schedule and the studio's cash projections were looking bad. The studio's cash statement and projection at this time is shown in Table 3.3. It indicated that the company would need to draw further on its overdraft facilities in the current quarter and would need more substantial funding in the new year. It was clear that even if all went well and they had no further problems, the soonest they could launch would be half-way through the following year. Even achieving this would probably involve what game developers call 'crunch': working

LightField Studios/Shutterstock

Table 3.3 Widescale's actual and projected cash flow (£000)

	Actual						Forecast					
Year	***1***		***2***				***3***				***4***	
Qtr	***3***	***4***	***1***	***2***	***3***	***4***	***1***	***2***	***3***	***4***	***1***	***2***
General revenue	2,273	2,332	2,105	2,117	2,306	,2308	2,205	1,886	2,000	2,000	2,000	2,000
General costs	1,891	1,764	1,792	1,898	1,894	1,869	1,800	1,800	1,800	1,800	1,800	1,800
Fierybryde revenue								0	800	4,500	3,500	2,500
Fierybryde costs	302	550	499	614	855	842	850	700	400	250	200	100
End of qtr. Cash	780	848	712	442	44	(386)	(831)	(1445)	(845)	3,605	7,105	9,705

extended hours of overtime (paid and unpaid) for periods of time in order to hit a particular deadline.

Oli was despondent: '*The real frustration is that the game is looking good. Everybody who has worked on it loves the storyline and are wowed by the graphics. We just need a final effort. It's tempting to see 'crunch' as a failure of planning. But honestly I've never worked on a development that has not involved some degree of crunch.*' It was Izzy who finally made the decision in the November. '*We have been working on this project for 18 months. That's not long for a game of this complexity. And it's really good, everyone agrees. But the potential of a game and the financial viability of its development process are different things. Basically, we have run out of credit and we have to accept that we need help. The most likely source of help is going to be a publisher. They could fund the remainder of this development from their small change. OK, they will demand a part of the company, and we would lose much of our independence. But it's either that, or abandon the development, let go probably a third of our staff, and try to get an emergency loan from our bank.*'

Questions

1. Was it a mistake for Widescale to embark on the Fierybryde development?
2. List the reasons that could have contributed to the Fierybryde development falling behind schedule.
3. What would you have done differently?
4. What would you advise Izzy and Oli to do now, and why?

Applying the principles

Some of these exercises can be answered by reading the chapter. Others will require some general knowledge of business activity and some might require an element of investigation. **Model answers for the first two questions are available on the student companion website.** *For model answers to the other questions in this section, please ask your tutor/lecturer.*

1. One product for which customers value a very wide range of product types is that of domestic paint. Most people like to express their creativity in their choice of paints and other home decorating products that they use in their homes. Clearly, offering a wide range of paint must have serious cost implications for the companies that manufacture, distribute and sell the product. Visit a store that sells paint and get an idea of the range of products available on the market. How do you think paint manufacturers and retailers could innovate to increase variety but minimise costs?

2. 'We have to get this new product out fast', said the Operations Director. 'Our competitors are close behind us and I believe their products will be almost as good as ours when they launch them.' She was talking about a new product that the company hoped would establish them as the leader in the market. The company had put together a special development team, together with their own development laboratory. They had spent £10,000 on equipping the laboratory and the cost of the development engineers would be £20,000 per quarter. It was expected that the new product would be fully developed and ready for launch within six quarters. It would be done through a specialist agency that charged £10,000 per quarter and would need to be in place two quarters prior to the launch. If the company met their launch date, it was expected that they could charge a premium price that would result in profits of approximately £50,000 per quarter. Any delay in the launch would result in a reduction in profits to £40,000 per quarter. If this development project were delayed by two quarters, how far would the break-even point for the project be pushed back?

3. Innovation becomes particularly important at the interface between offerings and the people that use them. Consider two types of website:

   a) those that are trying to sell something, such as Amazon.com; and
   b) those primarily concerned with giving information, for example reuters.com or nytimes.com.

   What constitutes good innovation for these two types of website? Find examples of particularly good and particularly poor web design and explain the issues you've considered in making the distinction between them.

4. According to the Ellen MacArthur Foundation, a circular economy is '*one that is restorative and regenerative by design, and which aims to keep products, components and materials at their highest utility and value at all times, distinguishing between technical and biological cycles*'. See also the example earlier in this chapter. What do you see as the main barriers to a more widespread adoption of the idea?

5. A janitor called Murray Spangler invented the vacuum cleaner in 1907. One year later he sold his patented idea to William Hoover, whose company went on to dominate the market. Now, the Dyson vacuum cleaner has jumped from nothing to a position where it dominates the market. The Dyson product dates back to 1978 when James Dyson designed a cyclone-based cleaner. It took five years and 5,000 prototypes before he had a working design. However, existing vacuum cleaner manufacturers were not as impressed – two rejected the design outright. So Dyson started making his new design himself and within a few years Dyson cleaners were outselling the rivals who had once rejected them. The aesthetics and functionality of the design help to keep sales growing in spite of a higher retail price. To Dyson good design '*is about looking at everyday things with new eyes and working out how they can be made better. It's about challenging existing technology*.'

   a) What was Spangler's mistake?
   b) What do you think makes 'good design' in markets such as the domestic appliance market?

c) Why do you think the two major vacuum cleaner manufacturers rejected Dyson's ideas?

d) How did design make Dyson a success?

6. It sounds like a joke, but it is a genuine product innovation. It's green, it's square and it comes originally from Japan. It's a square watermelon. Why square? Because Japanese grocery stores are not large and space cannot be wasted. Similarly, a round watermelon does not fit into a refrigerator very conveniently. There is also the problem of trying to cut the fruit when it keeps rolling around. An innovative Japanese farmer solved the problem with the idea of making a cube-shaped watermelon that could easily be packed and stored. There is no genetic modification or clever science involved in growing watermelons. It simply involves placing the young fruit into wooden boxes with clear sides. During its growth, the fruit naturally swells to fill the surrounding shape.

a) Why is a square watermelon an advantage?

b) What does this example tell us about product design?

Notes on chapter

1 The information on which this example is based is taken from: Economist (2018) 'The invention, slow adoption and near perfection of the zip', *Economist* print edition, 18 December.

2 Henderson, R.M. and Clark, K.B. (1990) 'Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms', *Administrative Science Quarterly*, 35 (1), pp. 9–30.

3 Economist (2019) 'Don't stop me now: The economics of streaming is changing pop songs', *Economist* print edition, 5 October.

4 The information on which this example is based is taken from: the Newlife Paints company website, www.newlifepaints.com [accessed 11 September 2020].

5 Anderson, E., Lin, S., Simester, D. and Tucker, C. (2015) 'Harbingers of failure', *Journal of Marketing Research*, 52 (5), pp. 580–592.

6 Christensen, C.M. (1997) *The Innovator's Dilemma: When new technologies cause great firms to fail*, Harvard Business Review Press.

7 Economist (2013) 'Be my guest: A short history of hotels', *Economist* print edition, 21 December.

8 The information on which this example is based is taken from: Economist (2017) 'One of the world's oldest products faces the digital future', *Economist* print edition, 12 October.

9 A term popularised largely by Chesbrough, for example see Chesbrough, H. and Garman, A. (2009) 'How open innovation can help you cope in lean times', *Harvard Business Review*, December; and Chesbrough, H. (2011) 'Bringing open innovation to services', *MIT Sloan Management Review*, 52 (2), pp. 85–90.

10 The information on which this example is based is taken from: BT website, 'How BT innovates', https://www.btplc.com/Innovation/HowBTinnovates/index.htm [accessed 11 September 2020]; BT News (2018) 'BT launches Better World Innovation Challenge for start-ups & SMEs', press release from BT, May 2018; BT Group plc Annual Report Strategic Report, 2019; Fransman, M. (2014) *Models of Innovation in Global ICT Firms: The emerging global innovation ecosystems*, JRC Scientific and Policy Reports – EUR 26774 EN, Seville: JRC-IPTS, https://ec.europa.eu/jrc/sites/jrcsh/files/jrc90726.pdf [accessed 29 September 2020].

11 Akella, R. (2018) 'What generative design is and why it's the future of manufacturing', newequipment.com, 16 March, https://www.newequipment.com/research-and-development/article/22059780/what-generative-design-is-and-why-its-the-future-of-manufacturing [accessed 11 September 2020].

12 Grieves, M.W. (2005) 'Product lifecycle management: The new paradigm for enterprises', *International Journal of Product Development*, 2 (1/2), pp. 71–84.

13 According to the TechJury website, the global games market for video games in 2019 was $148.8 billion: https://techjury.net/stats-about/video-games-industry/ [accessed 11 September 2020].

Taking it further

Chesbrough, H.W. ***(2003) Open Innovation: The new imperative for creating and profiting from technology, Harvard Business School Press.*** *A good overview of the reasons for open innovation and the challenges of making it work.*

Christensen, C.M. ***(1997) The Innovator's Dilemma: When new technologies cause great firms to fail, Harvard Business School Press.*** *Classic and ground-breaking book on disruptive innovation.*

Goffin, K. and Mitchell, R. ***(2016) Innovation Management: Effective strategy and implementation, 3rd edition, Palgrave Macmillan.*** *General advice from two experts in the subject.*

Gutsche, J. ***(2020) Create the Future + the Innovation Handbook: Tactics for disruptive thinking, Fast Company Press.*** *A popular and practical treatment; practitioners will gain much from it.*

Kahney, L. ***(2014) Jony Ive: The genius behind Apple's greatest products, Penguin.*** *Inside the mind of one of the world's great designers.*

Mckeown, M. ***(2008) The Truth About Innovation, Prentice Hall.*** *A practical guide to innovation, including lots of examples of success and failure.*

Nambisan, S. and Sawhney, M. ***(2007) 'A buyer's guide to the innovation bazaar', Harvard Business Review, June.*** *Provocative innovation ideas.*

Reason, B., Løvlie, L. and Brand Flu, M. ***(2016) Service Design for Business: A practical guide to optimizing the customer experience, John Wiley & Sons.*** *A book that looks at the 'entire service experience', and very readable.*

Tidd, J. and Bessant, J.R. ***(2018) Managing Innovation: Integrating technological, market and organizational change, 6th edition, John Wiley & Sons.*** *Well-established text that deserves its longevity.*

Trott, P. ***(2016) Innovation Management and New Product Development, 6th edition, Pearson.*** *A standard text in the subject area.*

4 Operations scope and structure

Introduction

Both the structure and the scope of an operation's supply network are decisions that shape how the operation interacts with its markets, its suppliers, in fact with the world in general. No operation exists in isolation. All operations are part of a larger and interconnected network of operations. This is called the operation's *supply network.* It will include the operation's suppliers and customers, as well as suppliers' suppliers and customers' customers, and so on. This chapter is concerned with the key strategic decisions that influence the scope and structure of this network, and more specifically the role of the individual operations within it. At a strategic level, operations managers are involved in deciding how much of the network an operation should own. This is called the *scope* of the operation. Put another way, the scope of the operation defines what it is going to do itself and what it will buy-in from suppliers. The managers are also concerned with the shape and form of their network. This is called the *structure* of the network. It involves deciding (or at least influencing) its overall shape, the location of each operation and how big the parts of the network that it owns should be. This chapter treats the issues related to both the scope and structure decisions (see Figure 4.1), and because both scope and structure decisions require an estimate of future demand, this chapter includes a supplement on forecasting.

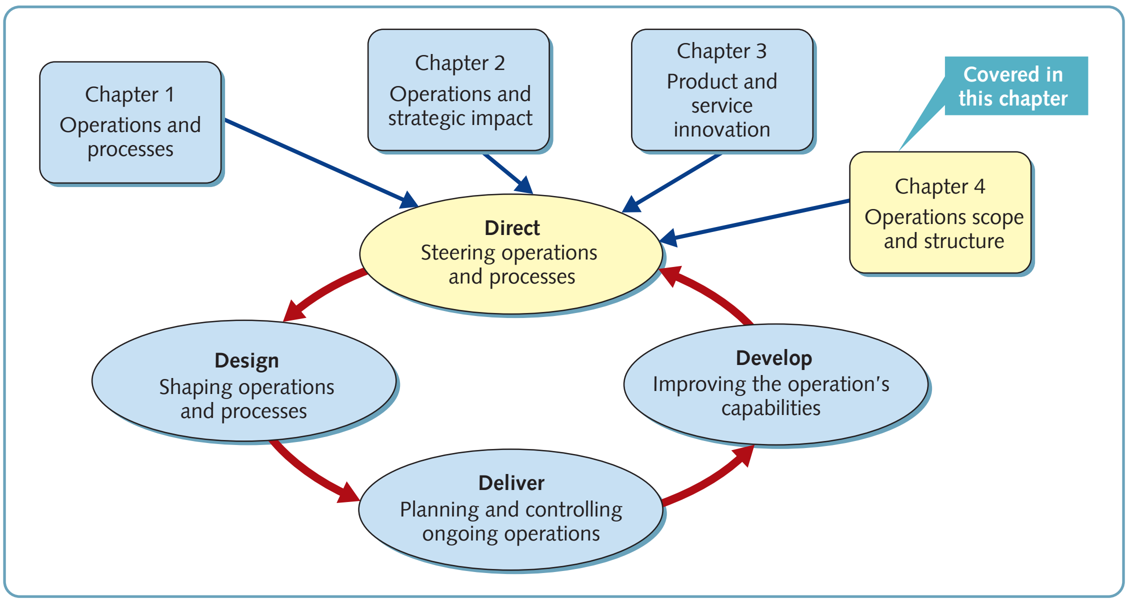
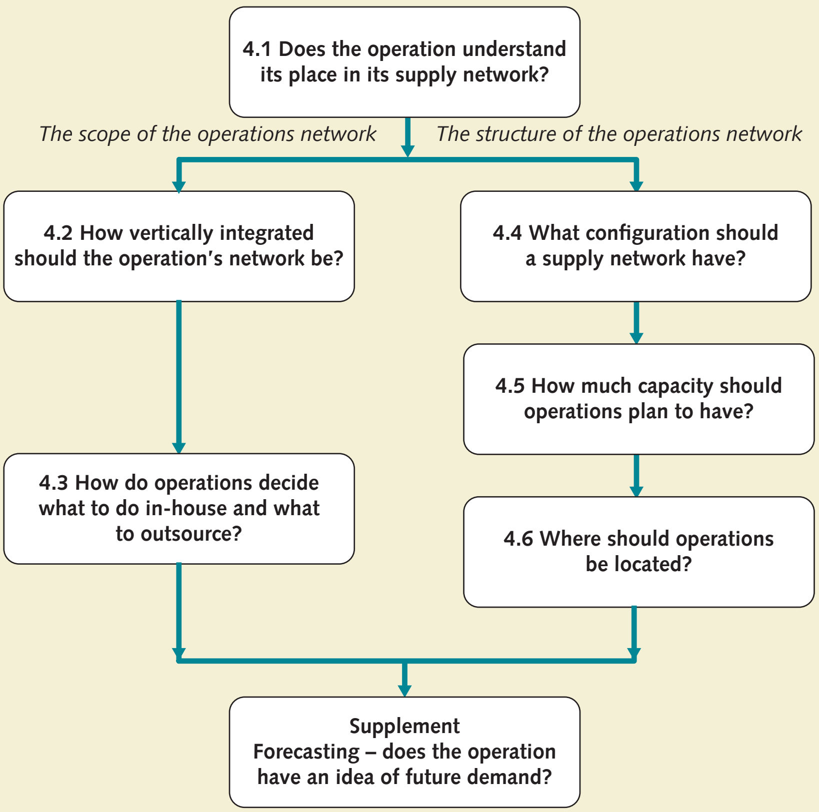


Figure 4.1 This chapter covers the structure and scope of operations

EXECUTIVE SUMMARY



4.1 Does the operation understand its place in its supply network?

The supply network includes the chains of suppliers providing inputs to the operation, the chain of customers who receive outputs from the operation, and sometimes other operations that may at times compete and other times cooperate. To understand its place and role, an operation must have an idea of how it fits into the scope and structure of its supply network. The 'structure' of an operation's supply network relates to the shape and form of the network. The 'scope' of an operation's supply network relates to the extent that an operation decides to perform the network activities itself, as opposed to requesting a supplier to do them. The structure and scope of an operation's supply network is important because it helps an understanding of competitiveness, it helps identify significant links in the network and it helps focus on long-term issues.

4.2 How vertically integrated should the operation's network be?

The scope to which an operation controls its supply network is the extent that it does things itself, as opposed to relying on other operations to do things for it. This is often referred to as 'vertical integration'. An organisation's vertical integration strategy can be defined in terms of the direction of integration, the extent of integration and the balance among the vertically integrated stages. The decision as to whether to vertically integrate is largely a matter of a business balancing the advantages and disadvantages as they apply to them.

4.3 How do operations decide what to do in-house and what to outsource?

Outsourcing is 'an arrangement in which one company provides services for another company that could also be, or usually have been, provided in-house'. The difference between vertical integration and outsourcing is largely a matter of scale and direction. Like the vertical integration decision, it is often a matter of balancing advantages against disadvantages under particular circumstances. Assessing the advisability of outsourcing should also include consideration of the strategic importance of the activity, the operation's relative performance and the risks of supply disruption.

4.4 What configuration should a supply network have?

Even when an operation does not directly own other operations in its network, it may still wish to change the shape of the network by reconfiguring it, in order to change the nature of the relationships. A number of trends are reshaping networks in many industries. These include reducing the number of individual suppliers, the disintermediation of some parts of the network and a greater tolerance of other operations being both competitors and complementors at different times (co-opetition). An idea that is closely related to that of co-opetition in supply networks is that of the 'business ecosystem', defined as 'an economic community supported by a foundation of interacting organisations and individuals'.

4.5 How much capacity should operations plan to have?

The amount of capacity an organisation will have depends on its view of current and future demand. It is when its view of future demand is different from current demand that this issue becomes important. When an organisation has to cope with changing demand, a number of capacity decisions need to be taken. These include choosing the optimum capacity for each site, balancing the various capacity levels of the operation in the network, and timing the changes in the capacity of each part of the network. Important influences on these decisions include the concepts of economy and diseconomy of scale.

4.6 Where should operations be located?

When operations change their location, their assumption is that the potential benefits of a new location will outweigh any cost and disruption involved in changing location. When operations do move, it is usually because of changes in demand and/or changes in supply. The factors that determine a location are such things as labour, land and utility costs, the image of the location, its convenience for customers and the suitability of the site itself.

4.1 Diagnostic question: *Does the operation understand its place in its supply network?*

The 'scope' of an operation's supply network relates to the extent that an operation decides to do the activities performed by the network itself, as opposed to requesting a supplier to do them. The 'structure' of an operation's supply network relates to the shape and form of the network. But before we examine these issues, we need to establish what we mean by a 'supply network' and why it is important that operations managers understand their position in it.

'*A supply network is an interconnection of organizations that relate to each other through upstream and downstream linkages between the different processes and activities that produce value in the form of products and services to the ultimate consumer*.'[1] In other words, a supply network is the means of setting an operation in the context of all the other operations with which it interacts, some of which are its suppliers and its customers. Materials, parts, other information, ideas and sometimes people all flow through the network of customer–supplier relationships formed by all these operations.

Terminology is important when describing supply networks. On its supply side an operation has its suppliers of parts, or information, or services. These suppliers themselves have their own suppliers who in turn could also have suppliers, and so on. On the demand side the operation has customers. These customers might not be the final consumers of the operation's products or services; they might have their own set of customers. Suppliers that supply the operation directly are often called first-tier suppliers. Second-tier suppliers supply them; however, some second-tier suppliers, as well as supplying first-tier suppliers, may also supply an operation directly, thus missing out a link in the network. Similarly, 'first-tier' customers are the main customer group for the operation. These in turn supply 'second-tier' customers, although again the operation may at times supply second-tier customers directly. The suppliers and customers that have direct contact with an operation are called its immediate supply network. Figure 4.2 illustrates this.

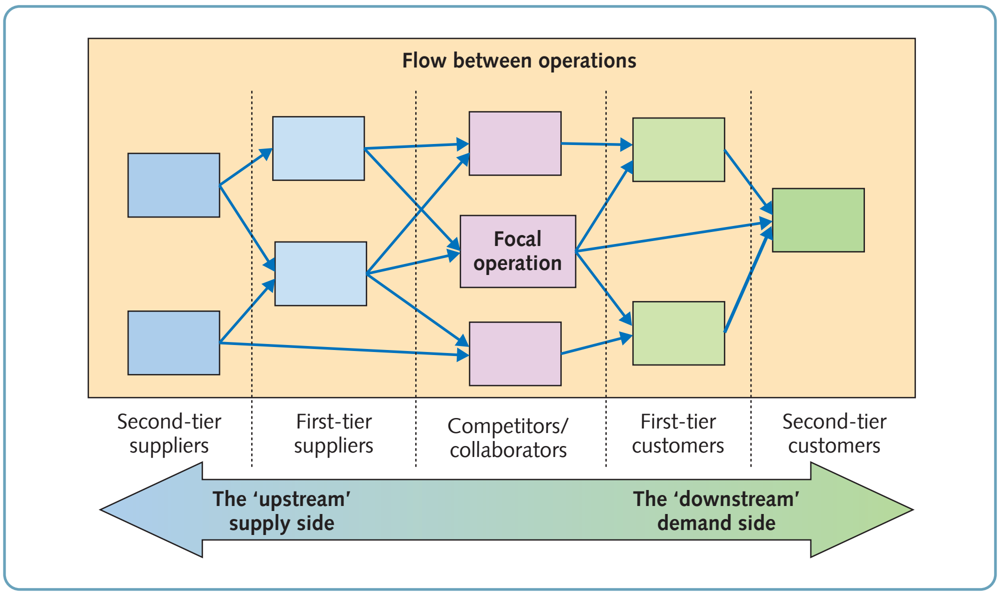


Figure 4.2 Supply network terminology

Two-way flow through the network

Materials, parts, information, ideas and sometimes people all flow through the network of customer–supplier relationships formed by all these operations. But also, along with the forward flow of transformed resources (materials, information and customers) in the network, each customer–supplier linkage will feed back orders and information. For example, when stocks run low, retailers place orders with distributors, who likewise place orders with the manufacturer; in turn the manufacturer will place orders with its suppliers, who will replenish their own stocks from their own suppliers. So flow is a two-way process with items flowing one way and information flowing the other.

Yet, it is not only manufacturers who are part of a supply network. The flow of physical materials may be easier to visualise, but service operations also have suppliers and customers who themselves have their own suppliers and customers. One way to visualise the supply networks of some service operations is to consider the downstream flow of information that passes between operations. Most financial service supply networks can be thought about like this. However, not all service supply networks deal primarily in information. For example, property companies that own and/or run shopping malls have suppliers who provide security services, cleaning services, maintenance services and so on. These first-tier suppliers will themselves receive service from recruitment agencies, consultants and so on. First-tier customers of the shopping mall are the retailers who lease retail space within the mall, who themselves serve retail customers. This is a supply network like any other. What is being exchanged between operations is the quality, speed, dependability, flexibility and cost of the services each operation supplies to its customers. In other words, there is a flow of 'operations performance' through the network. And although visualising the flow of 'performance' through supply networks is an abstract approach to visualising supply networks, it is a unifying concept. Broadly speaking, all types of supply network exist to facilitate the flow of 'operations performance'.

Case example

Virtually like Hollywood

Could that most ephemeral of all industries, Hollywood's film-making business, hold messages about scope and structure for even the most sober of operations? It is an industry whose complexity most of us do not fully appreciate. In American writer F. Scott Fitzgerald's unfinished novel *The Last Tycoon*, the narrator of the story, Cecelia Brady, said, '*You can take Hollywood for granted like I did, or you can dismiss it with the contempt we reserve for what we don't understand . . . not half a dozen men have ever been able to keep the whole equation of [making] pictures in their heads*'.[2] The 'equation' involves balancing the artistic creativity and fashion awareness necessary to create a market for its products, with the efficiency and tight operations practices that get films made and distributed on time. But although the form of the equation remains the same, the way its elements relate to each other has changed profoundly. The typical Hollywood studio once did everything itself. It employed everyone from the carpenters, who made the stage, through to the film stars. The film star Cary Grant (one of the biggest in his day) was as much of an employee as the chauffeur who drove him to the

Gabriele Maltinti/Shutterstock

studio, though his contract was probably more restrictive. The finished products were rolls of film that had to be mass produced and physically distributed to the cinemas of the world. No longer. Studios now deal almost exclusively in ideas. They buy and sell concepts, they arrange finance, they cut marketing deals and, above all, they manage the virtual network of creative and not-so-creative talent that goes into a film's production. The ability to put together teams of self-employed film stars and the small technical specialist operations that provide technical support has become a key skill. It is a world that is less easy for the studios, and even streaming services, to control.

Why is it important to consider the whole supply network?

What is undeniable is that supply network issues, and the operations' positions in them, have a significant effect on their strategic performance. In addition, there are other reasons why it is important to stand back and look at the whole (or a large part) of a supply network rather than an individual operation.

It helps an understanding of competitiveness

Immediate customers and immediate suppliers, quite understandably, are the main concern for companies. Yet sometimes they need to look beyond these immediate contacts to understand *why* customers and suppliers act as they do. Any operation has only two options if it wants to understand its ultimate customers' needs. It can rely on all the intermediate customers and customers' customers, and so on, that separate it from its end customers, or it can take responsibility for looking beyond its immediate customers and suppliers itself. Simply relying only on one's immediate network is arguably putting too much faith in someone else's judgement of things that are central to an organisation's own competitive health.

It helps identify significant links in the network

Not everyone in a supply network has the same degree of influence over the performance of the network as a whole. Some operations contribute more to the performance objectives that are valued by end customers. So an analysis of networks needs an understanding of the downstream and the upstream operations that contribute most to end-customer service. For example, the important end customers for domestic plumbing parts and appliances are the installers and service companies that deal directly with consumers. They are supplied by 'stock holders' who must have all parts in stock and deliver them fast. Suppliers of parts to the stock holders can best contribute to their end customers' competitiveness partly by offering a short delivery lead time but mainly through dependable delivery. The key players in this example are the stock holders. The best way of winning end-customer business in this case is to give the stock holder prompt delivery, which helps keep costs down while providing high availability of parts.

OPERATIONS PRINCIPLE
A supply network perspective helps to make sense of competitive, relationship and longer-term operations issues.

It helps focus on long-term issues

There are times when circumstances render parts of a supply network weaker than its adjacent links. High-street music stores, for example, have been largely displaced by music streaming and downloading services. A long-term supply network view would involve examining technology and market changes constantly to see how each operation in the supply network might be affected.

Scope and structure

The scope and structure of an operation's supply network are strongly related (which is why we treat them together). For example, suppose that a company that runs a shopping mall is dissatisfied with the service it is receiving from its supplier of security services. Also suppose that it is considering three alternatives. Option 1 is to switch suppliers and award the security contract to a different security services supplier. Option 2 is to accept an offer from the company that supplies cleaning services to supply both security and cleaning services. Option 3 is to take over responsibility for security itself, hiring its own security staff who would be put on the mall's payroll. These options are illustrated in Figure 4.3. The first of these options changes neither the structure nor the scope of this part of the supply network. The shopping mall still has three suppliers and is doing exactly what it did before. All that has changed is that security services are now being provided by another (hopefully better) supplier. However, option 2 changes the structure of the supply network (the mall now has only two suppliers, the combined cleaning and security supplier, and the maintenance supplier), but not the scope of what the mall does (it does exactly what it did before). Option 3 changes both the structure of the network (again, the mall has only two suppliers; cleaning and maintenance services) and the scope of what the mall does (it now also takes on responsibility for security itself).

A further point to make is that both scope and structure decisions actually comprise a number of other 'constituent' decisions.

The scope of an operation's activities within the network is determined by two decisions:

1. The extent and nature of the operation's vertical integration.
2. The nature and degree of outsourcing it engages in.

The structure of an operation's supply network is determined by three sets of decisions:

1. How the network should be configured.
2. The long-term capacity decision – what physical capacity each part of the network should have.
3. The location decision – where each part of the network should be located.

Note, however, that all of these decisions rely on forecasts of future demand, which the supplement to this chapter explores in more detail.

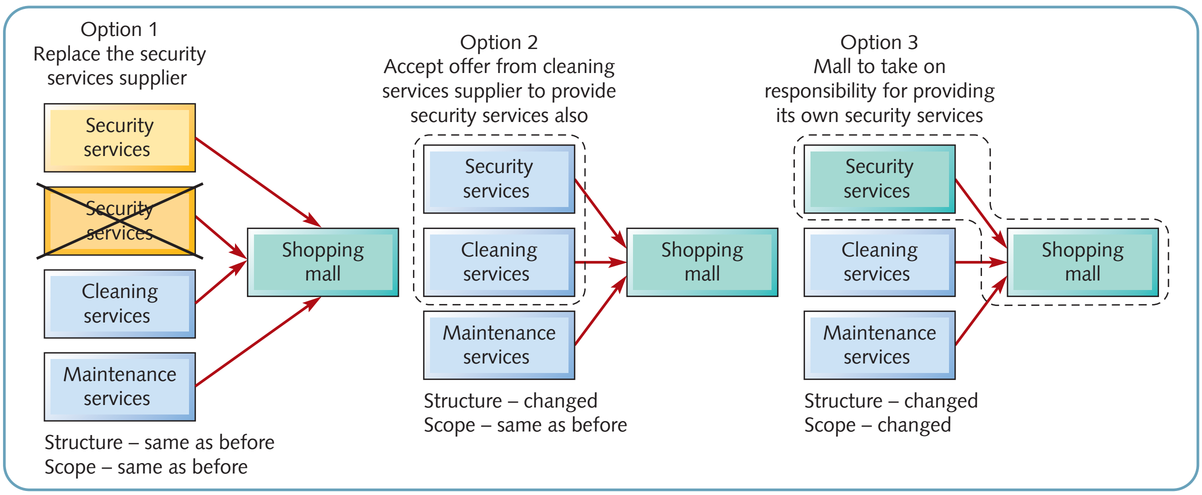


Figure 4.3 Three options for a shopping mall's supply network

Source: From Slack, N. *Operations Management*, 8e, © 2016 Pearson Education Limited, UK.

4.2 Diagnostic question: *How vertically integrated should the operation's network be?*

The scope of an operation's supply network determines to what extent an operation does things itself and the extent to which it will rely on other operations to do things for it. This is often referred to as 'vertical integration' when it is ownership of whole operations that are being decided, or 'outsourcing' when individual activities are being considered. We will look at the 'outsourcing' decision in the next section. Vertical integration is the extent to which an organisation owns the network of which it is a part. It usually involves an organisation assessing the wisdom of acquiring suppliers or customers. And different companies, even in the same industry, can make very different decisions over how much and where in the network they want to be. Figure 4.4 illustrates the (simplified) supply network for the wind turbine power generation industry. Original equipment manufacturers (OEMs) assemble the wind turbine nacelles (the nacelle houses the generator and gearbox). Towers and blades are often built to the OEM's specifications, either in-house or by outside suppliers. Installing wind turbines involves assembling the nacelle, tower and blades on site, erecting the tower and connecting to the electricity network. The extent of vertical integration varies by company and component. The three companies illustrated in Figure 4.4 have all chosen different vertical integration strategies. Company A is primarily a nacelle designer and manufacturer that also makes the parts. Company B is primarily an installer that also makes the tower and blades (but buys in the nacelle itself). Company C is primarily an operator that generates electricity and also designs and assembles the nacelles as well as installing the whole tower (but it outsources the manufacture of the nacelle parts, tower and blades).

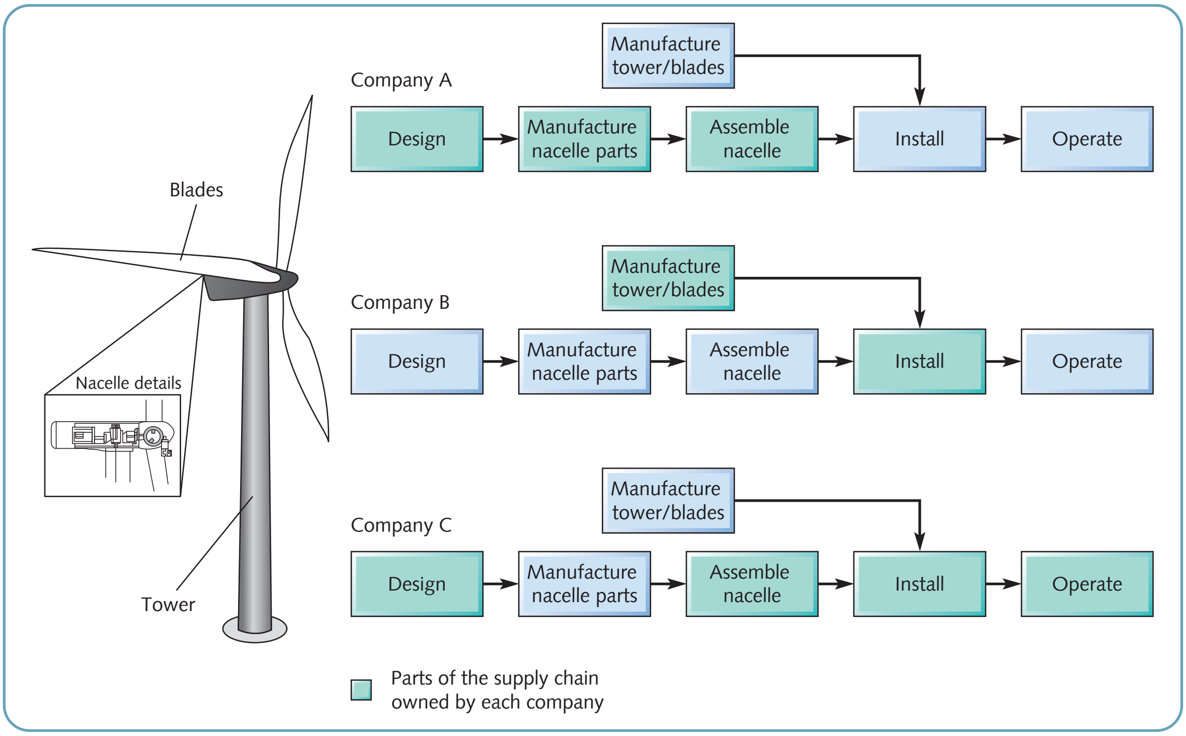


Figure 4.4 Three companies operating in the wind power generation industry with different vertical integration positions

Source: From Slack, N. *Operations Management*, *8e*, © 2016 Pearson Education Limited, UK.

Generally, an organisation's vertical integration strategy can be defined in terms of:

- *The direction of any integration* – does it expand by buying one of its suppliers (backward or 'upstream' vertical integration) or should it expand by buying one of its customers (forward or 'downstream' vertical integration)? Backward vertical integration, by allowing an organisation to take control of its suppliers, is often used either to gain cost advantages or to prevent competitors gaining control of important suppliers. Forward vertical integration, on the other hand, takes an organisation closer to its markets and allows more freedom for an organisation to make contact directly with its customers, and possibly sell complementary products and services.
- *The extent of the process span of integration* – some organisations deliberately choose not to integrate far, if at all, from their original part of the network. Alternatively, some organisations choose to become very vertically integrated.
- *The balance among the vertically integrated stages* – this is not strictly about the ownership of the network. It refers to the amount of capacity at each stage in the network that is devoted to supplying the next stage. So a totally balanced network relationship is one where one stage produces only for the next stage in the network and totally satisfies its requirements.

OPERATIONS PRINCIPLE
Vertical integration is the extent to which an organisation owns the network of which it is a part. It is defined by the direction, extent and balance of integration.

Advantages and disadvantages of vertical integration

The decision as to whether to vertically integrate in a particular set of circumstances is largely a matter of a business balancing the following advantages and disadvantages as they apply to them.

Perceived advantages of vertical integration

Although extensive vertical integration is no longer as popular as it once was, there are still companies who find it advantageous to own several sequential stages of their supply network. Indeed, very few companies are anywhere close to 'virtual', with no vertical integration of stages whatsoever. What then are the reasons why companies still choose to vertically integrate? Most justifications for vertical integration fall under four categories. These are:

1. ***It secures dependable access to supply or markets*** – the most fundamental reasons for engaging in some vertical integration are that it can give a more secure supply or bring a business closer to its customers. One reason why the oil companies, who sell gasoline, are also engaged in extracting it is to ensure long-term supply. In some cases, there may not even be sufficient capacity in the supply market to satisfy the company. It therefore has little alternative but to supply itself. Downstream vertical integration can give a firm greater control over its market positioning. For example, Apple has always adopted a supply network model where both its hardware and at least some of its software are 'designed' in-house, by Apple.
2. ***It may reduce costs*** – the most common argument here is that 'we can do it cheaper than our supplier's price'. Such statements are often made by comparing the marginal direct cost incurred by a company in doing something itself against the price it is paying to buy the product or service from a supplier. But costs savings should also take into account start-up and learning costs. A more straightforward case can be made when there are technical advantages of integration. For example, producing aluminium kitchen foil involves rolling it to the required thickness and then 'slitting' it into the finished widths. Performing both activities in-house saves the loading and unloading activity and the transportation to another operation. Vertical integration also reduces the 'transaction costs' of dealing with suppliers

and customers. Transaction costs are expenses, other than price, which are incurred in the process of buying and selling, such as searching for and selecting suppliers, setting up monitoring arrangements, negotiating contracts and so on. If transaction costs can be lowered to the point where the purchase price plus transaction costs is less than the internal cost, there is little justification for the vertical integration of the activity.

3. ***It may help to improve product or service quality*** – sometimes vertical integration can be used to secure specialist or technological advantage by preventing product and service knowledge getting into the hands of competitors. The exact specialist advantage may be anything from the 'secret ingredient' in fizzy drinks through to a complex technological process. In either case the argument is the same: 'this process gives us the key identifying factor for our products and services, so vertical integration therefore is necessary to the survival of product or service uniqueness'.

4. ***It helps in understanding other activities in the supply network*** – some companies, even those who are famous for their rejection of traditional vertical integration, do choose to own some parts of the supply network other than what they regard as core. For example, McDonald's, the restaurant chain, although largely franchising its retail operations, does own some retail outlets. How else, it argues, could it understand its retail operations so well?

Perceived disadvantages of vertical integration

The arguments against vertical integration tend to cluster around a number of observed disadvantages for those companies that have practised vertical integration extensively. These are:

1. ***It creates an internal monopoly*** – operations, it is argued, will only change when they see a pressing need to do so. Internal supply is less subject to the normal competitive forces that keep operations motivated to improve. If an external supplier serves its customers well, it will make higher profits; if not, it will suffer. Such incentives and sanctions do not apply to the same extent if the supplying operation is part of the same company.

2. ***You can't exploit economies of scale*** – any activity that is vertically integrated within an organisation is probably also carried out elsewhere in the industry. But the effort it puts into the process will be a relatively small part of the sum total of that activity within the industry. Specialist suppliers who can serve more than one customer are likely to have volumes larger than any of their customers could achieve by doing things for themselves. This allows specialist suppliers to reap some of the cost benefits of economies of scale, which can be passed on in terms of lower prices to their customers.

3. ***It results in loss of flexibility*** – heavily vertically integrated companies, by definition, do most things themselves. This means that a high proportion of their costs will be fixed costs. They have, after all, invested heavily in the capacity that allows them to do most things in-house. A high level of fixed costs relative to variable costs means that any reduction in the total volume of activity can easily move the economics of the operation close to, or below, its break-even point.

4. ***It cuts you off from innovation*** – vertical integration means investing in the processes and technologies necessary to produce products and services in-house. But, as soon as that investment is made, the company has an inherent interest in maintaining it. Abandoning such investments can be both economically and emotionally difficult. The temptation is always to wait until any new technology is clearly established before admitting that one's own is obsolete. This may lead to a tendency to lag in the adoption of new technologies and ideas.

5. ***It distracts you from core activities (loss of focus)*** – the final, and arguably most powerful, case against vertical integration concerns any organisation's ability to be technically competent at a very wide range of activities. All companies have things that they need to be good at. And it is far easier to be exceptionally good at something if the company focuses exclusively on it, rather than being distracted by many other things. Vertical integration, by definition, means doing more things, which can distract from the (few) particularly important things.

4.3 Diagnostic question: *How do operations decide what to do in-house and what to outsource?*

Theoretically, the 'vertical integration' decision and the 'outsourcing' decision are almost the same thing. The difference between them is one of scale. Vertical integration is a term that is usually (but not always) applied to whole operations. So, buying a supplier because you want to deny their products to a competitor, or selling the part of your business that services your products to a specialist servicing company that can do the job better, are vertical integration decisions. Outsourcing usually applies to smaller sets of activities that have previously been performed in-house. Deciding to ask a specialist laboratory to perform some quality tests that your own quality control department used to do, or having your call (contact) centre taken over and run by a larger call-centre company, are both outsourcing decisions.

Outsourcing is also known as the 'do-or-buy' decision, and has become an important issue for most businesses. This is because, although most companies have always outsourced some of their activities, a larger proportion of direct activities are now bought from suppliers. Also, many indirect and administrative processes are now outsourced. This is often referred to as business process outsourcing (BPO). Financial service companies, in particular, outsource some of their more routine back-office processes. In a similar way, many processes within the human resource function, from simple payroll services through to more complex training and development processes, are outsourced to specialist companies. The processes may still be physically located where they were before, but the outsourcing service provider manages the staff and technology. The reason for doing this is often primarily to reduce costs. However, there can sometimes also be significant gains in the quality and flexibility of service offered.

Making the outsourcing decision

Outsourcing is rarely a simple decision. Operations in different circumstances with different objectives are likely to make different decisions. Yet the question itself is relatively simple, even if the decision is not: 'Is it in-house or outsourced supply in a particular set of circumstances that gives the company the appropriate performance objectives that it requires to compete more effectively in its markets?' For example, if the main performance objectives for an operation are dependable delivery and meeting short-term changes in customers' delivery requirements, the key question should be: 'How does in-house or outsourcing give better dependability and delivery flexibility performance?' This means judging two sets of opposing factors – those that give the potential to improve performance, and those that work against this potential being realised. Table 4.1 summarises some arguments for in-house supply versus outsourcing in terms of each performance objective.

> **OPERATIONS PRINCIPLE**
> ***Assessing the advisability of outsourcing should include how it impacts on relevant performance objectives.***

Table 4.1 How in-house and outsourced supply may affect an operation's performance objectives

Performance objective	*'Do it yourself' in-house supply*	*'Buy it in' outsourced supply*
Quality	The origins of any quality problems are usually easier to trace in-house and improvement can be more immediate, but there can be some risk of complacency.	Supplier may have specialised knowledge and more experience, also may be motivated through market pressures, but communication more difficult.
Speed	Can mean synchronised schedules, which speeds throughput of materials and information, but if the operation has external customers, internal customers may be low priority.	Speed of response can be built into the supply contract where commercial pressures will encourage good performance, but there may be significant transport/delivery delays.
Dependability	Easier communications can help dependability but, if the operation also has external customers, internal customers may receive low priority.	Late delivery penalties in the supply contract can encourage good delivery performance, but organisational barriers may inhibit communication.
Flexibility	Closeness to the real needs of a business can alert the in-house operation to required changes, but the ability to respond may be limited by the scale and scope of internal operations.	Outsourced suppliers may be larger with wider capabilities than in-house suppliers and have more ability to respond to changes, but may have to balance conflicting needs of different customers.
Cost	In-house operations do not have to make the margin required by outside suppliers, so the business can capture the profits that would otherwise be given to the supplier, but relatively low volumes may mean that it is difficult to gain economies of scale or the benefits of process innovation.	Probably the main reason why outsourcing is so popular, as outsourced companies can achieve economies of scale and they are motivated to reduce their own costs because it directly impacts on their profits, but costs of communication and coordination with supplier need to be taken into account.

Case example

Compass and Vodafone – two ends of the outsourcing phenomenon[3]

Some companies have built their success on offering outsourcing services. Others are the customers of outsourcing companies. Few are neither; both private and public organisations have outsourced an increasing proportion of their activities over the last few decades. The most common type of outsourcing is when an operation delegates the responsibility for non-core, but also non-trivial, activities to an external specialist service supplier. For most organisations, an obvious candidate for outsourcing is their in-house catering activities. Which is why there are many food service companies willing to take on that responsibility. The biggest in the world is the Compass Group, which provides service for a wide range of customers, from factories and office canteens, to schools and universities, oil rigs, the military, prisons and sporting events. Business customers that have outsourced their catering to Compass include Google, Microsoft, Nike, HSBC and Intel. Serving five and a half *billion* meals every year in 55,000 client locations, its 600,000 employees meet the catering needs of many different types of client. The group has gained its leading position in foodservice outsourcing partly because of its sector approach, which distinguishes between the varying market needs of 'business and industry', 'healthcare and seniors', 'education', 'sports and leisure' and 'defence, offshore and remote'. But, although Compass has been particularly successful in establishing itself as a leading player in the industry, there are very good reasons for any organisation to consider outsourcing its catering operations. For most businesses, catering is clearly not their core business. However good an in-house operation is, it will not be able to provide the variety or have the expertise or the scale of a company like Compass. Nor is it likely to be able to stay ahead of food trends. Also, a company like Compass could help in providing non-routine catering (such as events), reception services, office cleaning and some facilities management.

Any organisation that gets its outsourced catering wrong runs the risk of upsetting its staff. Any organisation that gets its outsourced customer servicing wrong runs the risk of upsetting its customers. Which, in turn, impacts its revenues, and its future. But the arguments for outsourcing customer servicing are similar to outsourcing catering. It holds the potential to give the same quality of customer support as in-house operations, but at a lower cost, particularly if it is outsourced to a country with a lower cost of living. It could provide a wider range of services and meet unexpectedly high demand. And specialist outsource operations can invest in the latest technologies. Even so, customer care can be seen as so important that outsourcing it will always involve some risk. Customer frustration, especially if routed to overseas locations, can prompt companies to consider bringing call-centre operations back to their home countries. One company to do this is Vodaphone, the UK-based telecommunications company, which announced that it would be transferring over 2,000 jobs from its overseas call centres (mainly in South Africa, where it was using an external agency) to the United Kingdom. The new roles, it said, would be spread across the existing Vodafone call centres in various parts of the UK. They said that reshoring of the call-centre jobs would bring the company's consumer mobile customer services in line with its UK-based call centres for UK business and broadband customers, and that the new roles would make a real difference to its customers as well as a real difference to the communities that would be the focus of its customer services investment. However, although Vodaphone's move was directed at enhancing the quality of its UK customer services operations, in some parts of the world legislation can be a factor. For example, the Italian government introduced a law that gave consumers calling a company the option of speaking to a call-centre worker in Italy rather than someone overseas.

Incorporating strategic factors into the outsourcing decision

Although the effect of outsourcing on the operation's performance objectives is important, there are other factors that companies take into account when deciding if outsourcing an activity is a sensible option. For example, if an activity has long-term strategic importance to a company, it is unlikely to outsource it. A retailer might choose to keep the design and development of its website in-house for example, because it plans to move into web-based retailing at some point in the future, even though specialists could perform the activity at less cost. Nor would a company usually outsource an activity for which it had specialised skills or knowledge. For example, a company making laser printers may have built up specialised knowledge in the production of sophisticated laser drives. This capability may allow it to introduce product or process innovations in the future. It would be foolish to 'give away' such capability. After these two more strategic factors have been considered, the company's operations performance can be taken into account. Obviously if its operations performance is already far superior to any potential supplier, it would be unlikely to outsource the activity. But even if its performance is currently below that of potential suppliers, it may not outsource the activity if it feels that it could significantly improve its performance. Figure 4.5 illustrates this decision logic.

OPERATIONS PRINCIPLE
Assessing the advisability of outsourcing should include consideration of the strategic importance of the activity and the relative operations performance.

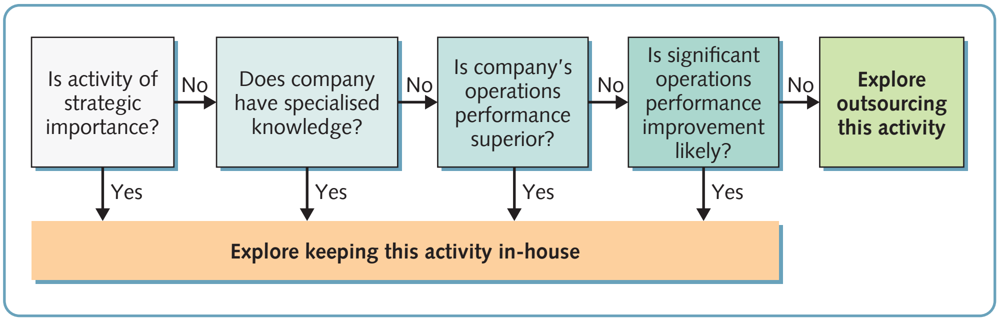


Figure 4.5 The decision logic of whether to outsource or not

Outsourcing and offshoring

Two supply network strategies that are often confused are those of outsourcing and offshoring. Outsourcing means deciding to buy-in products or services rather than perform the activities in-house. Offshoring means obtaining products and services from operations that are based outside one's own country. Of course, one may both outsource and offshore. Offshoring is very closely related to outsourcing and the motives for each may be similar. Offshoring to a lower-cost region of the world is usually done to reduce an operation's overall costs, as is outsourcing to a supplier who has greater expertise or scale, or both.

Globalisation and 'reshoring'

Almost since the start of the Industrial Revolution, businesses across the world have been forging ever-closer relationships. Especially in the last few decades, the use of geographically dispersed suppliers to outsource at least some activities has become routine. This has been driven partly by labour cost differentials, partly by cheap and efficient telecommunication between businesses, partly by trade deals and partly by reducing transport costs. This is 'globalisation', where products, raw materials, money, technology and ideas move (relatively) smoothly across national boundaries. Apple and others can design their products in California, while Asian assembly lines assemble them. A French aerospace company can direct the activities of its Brazilian suppliers almost as effectively as if they were in the next town. Bill Clinton, the ex-US President, called globalisation 'the economic equivalent of a force of nature'.

But increasingly some economists and business commentators question whether the boom in globalised operations is over. Some cite protectionist pressures in some more economically developed countries. Others see rising wages in (previously) less-developed countries as reducing cost differentials. In addition, the operations-related advantages of sourcing from nearby suppliers can be significant. Reducing reliance on complicated international supply chains can save transport and inventory costs, and is less polluting and potentially less prone to reputational risk if far-off suppliers misbehave. It also could increase supply flexibility. For example, the Spanish fast-fashion brand Zara manufactures some of its 'steady-selling' items in low-cost factories in Asia, but makes those garments with less predictable demand closer to its markets so that it can respond quickly to changing fashions. Developments in technology could reinforce this so-called 'reshoring' process. Automation may encourage a trend towards 'radical insourcing', where more economically developed countries no longer need to outsource production to countries where wages are low.

Risk and reshoring

In the case example, 'Compass and Vodaphone – two ends of the outsourcing phenomenon', the main reason that Vodaphone reshored some of its call-centre capacity was to increase the perceived service that customers received. Since the COVID-19 pandemic, the risks of depending on offshore supply has also become something that is increasingly considered. While demand for some 'non-essential' goods almost dried up (like clothing), other items that almost all countries wanted were suddenly in desperately short supply. (In the international scramble, medical supplies from face shields to nose swabs became rare items.) It was not just that supply chains were disrupted, that had happened before. It was that all parts of the world were affected. For many operations there was no obvious short-term way out of their supply problems. Nor would any alternative supply arrangement have saved them. Even bringing supply closer might not have helped particularly. Nevertheless, shorter supply chains and ultimately bringing 'supply' in-house, does give a sense of more control. Which is why some firms, for the first time, seriously started to consider bringing at least some supplies back closer to home, or even back in-house.

4.4 Diagnostic question: *What configuration should a supply network have?*

'Configuring' a supply network means determining its overall pattern. In other words, what should be the pattern, shape or arrangement of the various operations that make up the supply network? Even when an operation does not directly own, or even control, other operations in its network, it may still wish to change the shape of the network. This involves attempting to manage network behaviour by reconfiguring the network to change the nature of the relationships between its various parts. Reconfiguring a supply network sometimes involves parts of the operation being merged – not necessarily in the sense of a change of ownership of any parts of an operation, but rather in the way responsibility is allocated for carrying out activities. The most common example of network reconfiguration has come through the many companies that have recently reduced the number of their direct suppliers. The complexity of dealing with many hundreds of suppliers may both be expensive for an operation and (sometimes more importantly) prevent the operation from developing a close relationship with a supplier. It is not easy to be close to hundreds of different suppliers.

Disintermediation

Another trend in some supply networks is that of companies within a network bypassing customers or suppliers to make contact directly with customers' customers or suppliers' suppliers. 'Cutting out the middle men' in this way is called disintermediation. An obvious example of this is the way the internet has allowed some suppliers to 'disintermediate' traditional retailers in supplying goods and services to consumers. So, for example, many services in the travel industry that used to be sold through retail outlets (travel agents) are now also available direct from the suppliers. The option of purchasing the individual components of a vacation through the websites of the airline, hotel, car-hire operation and so on, is now easier for consumers. Of course, they may still wish to purchase an 'assembled' product from retail travel agents, which can have the advantage of convenience. Nevertheless, the process of disintermediation has developed new linkages in the supply network.

Co-opetition

One approach to thinking about supply networks sees any business as being surrounded by four types of players: suppliers, customers, competitors and complementors. Complementors enable one's products or services to be valued more by customers because they can also have the complementor's products or services, as opposed to when they have yours alone. Competitors are the opposite; they make customers value your product or service less when they can have the competitor's product or service, rather than yours alone. Competitors can also be complementors and vice versa. For example, adjacent restaurants may see themselves as competitors for customers' business. A customer standing outside and wanting a meal will choose between the two of them. Yet, in another way they are complementors. Would that customer have come to this part of town unless there was more than one restaurant to choose from? Restaurants, theatres, art galleries and tourist attractions generally all cluster together in a form of cooperation to increase the total size of their joint market. It is important to distinguish between the way companies cooperate in increasing the total size of a market and the way in which they then compete for a share of that market. Customers and suppliers, it is argued, should have 'symmetric' roles. Harnessing the value of suppliers is just as important as listening to the needs of customers. Destroying value in a supplier in order to create it in a customer does not increase the value of the network as a whole. So, pressurising suppliers will not necessarily add value. In the long term it creates value for the total network to find ways of increasing value for suppliers as well as customers. All the players in the network, whether they are customers, suppliers, competitors or complementors, can be both friends and enemies at different times. The term used to capture this idea is 'co-opetition'.

The idea of the 'business ecosystem'[4]

An idea that is closely related to that of co-opetition in supply networks is that of the 'business ecosystem'. It can be defined as: '*An economic community supported by a foundation of interacting organizations and individuals – the organisms of the business world. The economic community produces goods and services of value to customers, who are themselves members of the ecosystem. The member organisms also include suppliers, lead producers, competitors, and other stakeholders. Over time, they coevolve their capabilities and roles, and tend to align themselves with the directions set by one or more central companies.*'[5] One of the main differences between this idea and that of the supply network generally is the inclusion in the idea of the ecosystem of businesses that may have little or no direct relationship with the main supply network, yet exist only because of that network. They interact with each other, predominantly complementing or contributing significant components of the value proposition for customers. Many examples come from the technology industries. The innovative products and services that are developed in the technology sectors cannot evolve in a vacuum. They need to attract a whole range of resources, drawing in expertise, capital, suppliers and customers to create cooperative networks. For example, the app developers that develop applications for particular operating system platforms may not be 'suppliers', as such, but the relationship between them and the supply network that supplies the mobile device is mutually beneficial. Building an ecosystem of developers around a core product can increase its value to the end customer and by doing so complements the usage of the core product. Such an ecosystem of complementary products and services can also create significant barriers to entry for new competitors. Any possible competitors would not only have to compete with the core product, but they would also have to compete against the entire ecosystem of complementary products and services.

Describing supply networks – dyads and triads

The supply network that was illustrated in Figure 4.2 is, of course, a simplification. Any realistic supply network diagram will be much more complex. There are many operations, all interacting in different ways to produce end products and services. Because of this, and to understand them better, supply network academics and professionals often choose to focus on the individual interaction between two specific operations in the network. This is called a 'dyadic' (simply meaning 'two') interaction, or dyadic relationship, and the two operations are referred to as a 'dyad'. So, if one wanted to examine the interactions that a focal operation had with one of its suppliers and one of its customers, one would examine the two dyads of 'supplier–focal operation', and 'focal operation–customer', see Figure 4.6(a). For many years most discussion (and research) on supply networks was based on dyadic relationships. This is not surprising as all relationships in a network are based on the simple dyad. However, more recently, and certainly when examining service supply networks, many authorities make the point that dyads do not reflect the real essence of a supply network. Rather, they say, it is triads, not dyads, that are the basic elements of a supply network, see Figure 4.6(b). No matter how complex a network, it can be broken down into a collection of triadic interactions. The idea of triads is especially relevant in service supply networks. Operations are increasingly outsourcing the delivery of some aspects of their service to specialist providers, who deal directly with customers on behalf of the focal operation (more usually called the 'buying operation', or just 'buyer' in this context). For example, Figure 4.6(b) illustrates the common example of an airline contracting a specialist baggage-handling company to provide services to its customers

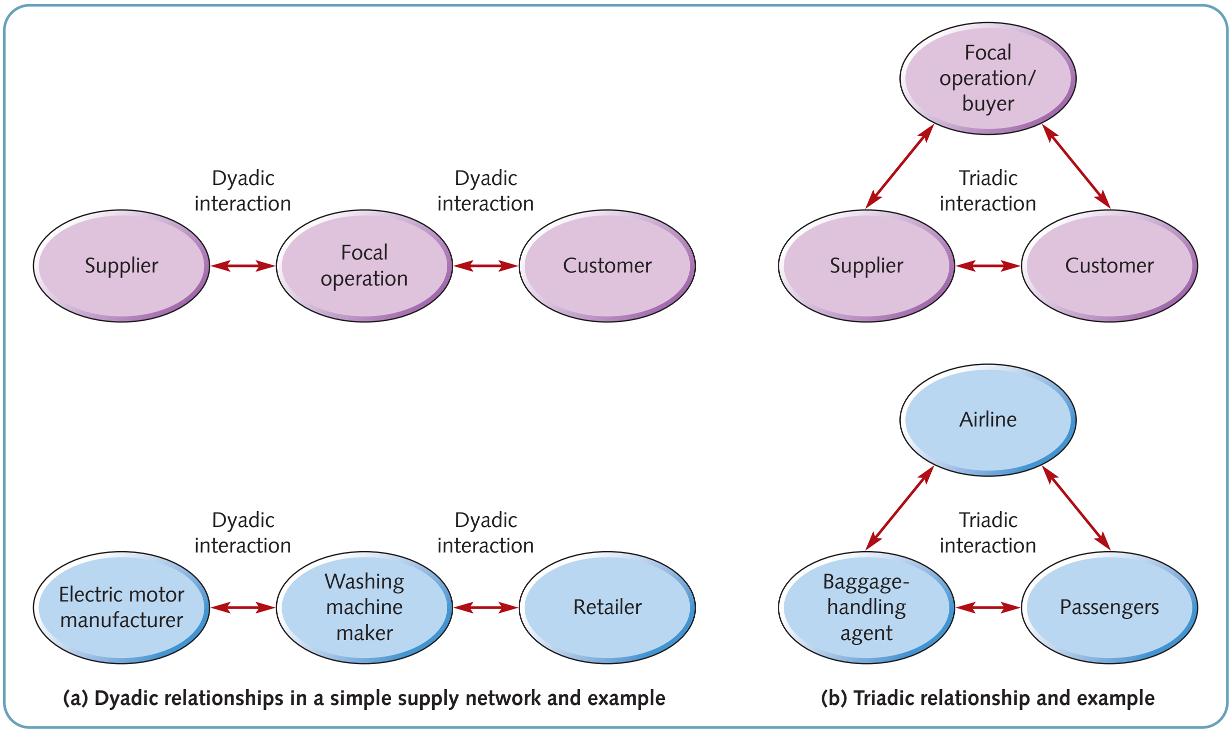


Figure 4.6 Dyadic and triadic relationships in two simple supply networks and examples
Source: From Slack, N. *Operations Management*, 8e, © 2016 Pearson Education Limited, UK.

on its behalf. Similarly, internal services are increasingly outsourced to form internal triadic relationships. For example, if a company outsources its IT operations, it is forming a triad between whoever is purchasing the service on behalf of the company, the IT service provider and the employees who use the IT services.

Thinking about supply networks as a collection of triads rather than dyads is strategically important. First, it emphasises the dependence that organisations place on their suppliers' performance when they outsource service delivery. A supplier's service performance makes up an important part of how the buyer's performance is viewed. Second, the control that the buyer of the service has over service delivery to its customer is diminished in a triadic relationship. In a conventional supply chain, with a series of dyadic relationships, there is the opportunity to intervene before the customer receives the product or service. However, products or services in triadic relationships bypass the buying organisation and go directly from provider to customer. Third, and partially as a consequence of the previous point, in triadic relationships the direct link between service provider and customer can result in power gradually transferring over time from the buying organisation to the supplier that provides the service. Fourth, it becomes increasingly difficult for the buying organisation to understand what is happening between the supplier and customer at a day-to-day level. It may not even be in the supplier's interests to be totally honest in giving performance feedback to the buyer. Finally, this closeness between supplier and customer, if it excludes the buyer, could prevent the buyer from building important knowledge. For example, suppose a specialist equipment manufacturer has outsourced the maintenance of its equipment to a specialist provider of maintenance services. The ability of the equipment manufacturer to understand how its customers are using the equipment, how the equipment is performing under various conditions, how customers would like to see the equipment improved, and so on, is lost. The equipment manufacturer may have outsourced the cost and trouble of providing maintenance services, but it has also outsourced the benefits and learning that come from direct interaction with customers.

Structural complexity in supply networks

Some supply networks are relatively straightforward, both to describe and to manage. Others are more complex. At a simple level, the three-stage dyadic supply relationship shown in Figure 4.6(a) is less complex than the triadic supply network relationship shown in Figure 4.6(b). Put many triadic relationships together into an interrelated network and things become even more complex. This is especially true when customers have some degree of choice over how and when they interact with a network. An example of such structural complexity is the move towards what have been called 'omnichannel retail networks'. The original relationship between a retailer and its customers was straightforward – the retailer expected customers to collect goods from its stores, or delivered them later if requested by the customer. When online retailers emerged as the internet developed, there was little overlap between conventional 'bricks-and-mortar' stores and online firms. Even when conventional retailers developed an online presence, they often kept their online operations separate from their high-street stores.This is the 'single-channel' model, as shown in Figure 4.7.

As more methods of contacting retailers became available, such as mobile phones, apps and social media, most retailers struggled to manage and fully integrate the many

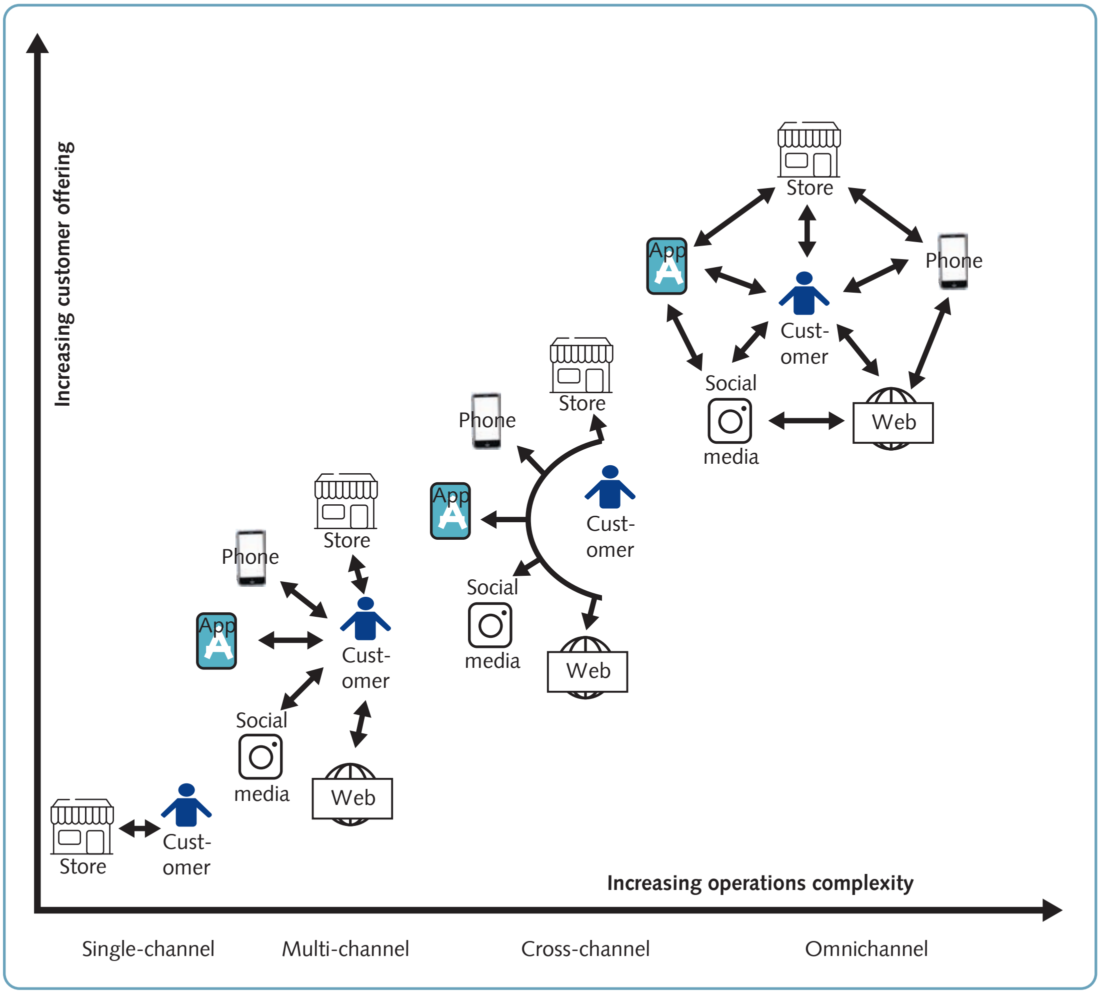


Figure 4.7 Single-channel, multi-channel, cross-channel and omnichannel retail models represent increasing options for customers, but increasing operations complexity

alternative 'channels' of communication. In fact, they often treated them specifically as independent entities so that they could align each channel with specific targeted customer segments. This is the 'multi-channel' model shown in Figure 4.7. Later came the first attempts to integrate high-street stores with (initially) web and other channels in order to enhance their cross-functionality that, in turn, allowed a better shopping experience for customers (see the 'cross-channel' element in Figure 4.7). Finally, the 'omnichannel' model (also depicted in Figure 4.7) is seen by many retail analysts as being one of the most important retail developments since the advent of online services. It seeks to provide a seamless all-inclusive customer experience by fully integrating all possible channels, so customers can use whichever is the most convenient for them at whatever stage of the transaction. So, a customer could browse alternative products through social media, order their choice via an app, manage their account through a website, pay using their phone and return the product to a physical store, should they wish to. Obviously, this requires a degree of technical sophistication as well as coordination between internal functions such as marketing, retailing operations, distribution and IT.

Case example

Amazon and Bonobos move to include bricks and mortar[6]

The customer-facing part of supply networks – that is, the bit that delivers to the end customer – has been subject to a huge amount of disruption since the rise of online retailers. Almost all types of consumer products, from groceries to garments, had well-established online retailers battling with their 'bricks-and-mortar' competitors. Then the distinction between online and physical 'high-street' retailers started to become less clear. The first move was for traditional high-street retail brands to develop an online presence. Those who did not do this often suffered compared with those that successfully integrated their online and high-street offerings. Then, rather to the surprise of some experts, some online retailers started to open physical shops. Some of the first were sellers of 'big ticket' items like furniture. Many customers liked to see and examine their sofa before ordering it. Then the idea spread to retailers of more portable items. For example, Amazon, the world's largest e-commerce company, has begun to open physical 'bricks-and-mortar' retail operations. Its first move was physical bookstores, after which it acquired Whole Foods Market, an upmarket organic supermarket chain. It has also employed identification technology to open Amazon Go, its cashier-less grocery stores. Similarly, Bonobos, the American men's garment retailer, owned by Walmart, started as an exclusively online operation, but decided to open physical stores. Yet, like some furniture showrooms, but unlike most garment retailers, Bonobos shops did not hold stock for sale to its customers. The idea was to separate the purchase of a product from its distribution. This meant that customers were able to 'get the feel' of its clothing, without the costs of a conventional store. Storing and selling goods in the same location used extra (and often expensive) space, and distracted employees who had to spend time managing the inventory and restocking shelves, rather than attending to customers. Sales staff could focus on helping their customers to find clothes, ascertain the proper fit and order the clothes online. Like other retailers that followed them, Bonobos found that they could do a better job of selling clothes because they did not stock the clothes.

AP/Shutterstock

4.5 Diagnostic question: *How much capacity should operations plan to have?*

The next set of 'structure' decisions concern the size or capacity of each part of the supply network. Here we shall treat capacity in a general long-term sense. The specific issues involved in measuring and adjusting capacity in the medium and short terms are examined in Chapter 8.

The optimum capacity level

Most organisations have decisions to make about how big (in terms of capacity) they want to be. A chain of truck service centres, for example, might operate centres that have various capacities. The effective cost of running each centre will depend on the average service-bay occupancy. Low occupancy, because of few customers, will result in a high cost per customer served because the fixed costs of the operation are being shared between few customers. As demand, and therefore service-bay occupancy, increases, the cost per customer will reduce. However, operating at very high levels of capacity utilisation (occupancy levels close to capacity) can mean longer customer waiting times and reduced customer service. There may also be less-obvious penalties for operating centres at levels close to their nominal capacity. For example, long periods of overtime may reduce productivity levels, as well as costing more in extra payments to staff; utilising bays at very high levels reduces the time available for maintenance and cleaning, which may increase the number of breakdowns, reduce effective life, and so on. This usually means that average costs start to increase after a point that will often be lower than the theoretical capacity of the operation.

> OPERATIONS PRINCIPLE
> *All types of operations exhibit economy-of-scale effects where operating costs reduce as the scale of capacity increases.*

The blue curves in Figure 4.8 show this effect for the service centres of 5-, 10- and 15-bay capacity. As the nominal capacity of the centres increases, the lowest cost point at first reduces. This is because the fixed costs of any operation do not increase proportionately as its capacity increases. A 10-bay centre has less than twice the fixed costs of a 5-bay centre. The capital costs of constructing the operations do not increase proportionately to their capacity. A 10-bay centre costs less to build than twice the cost of building a 5-bay centre. These two factors, taken together, are often referred to as economies of scale – a universal concept that applies (up to a point) to all types of operation. However, economies of scale do not go on forever. Above a certain size, the lowest cost point on curves such as that shown in Figure 4.8 may increase. This occurs because of what are called diseconomies of scale, two of which are

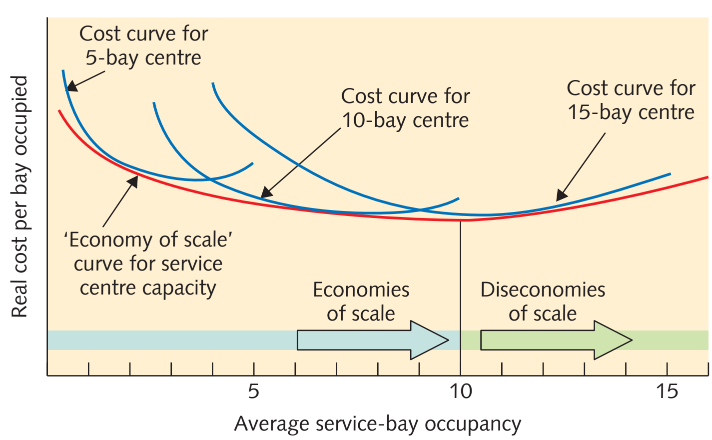


Figure 4.8 Unit-cost curves for individual service centres of varying capacities

particularly important. First, complexity costs increase as size increases. The communications and coordination effort necessary to manage an operation tends to increase faster than capacity. Although not seen as a direct cost, this can nevertheless be very significant. Second, a larger centre is more likely to be partially underutilised because demand within a fixed location will be limited. The equivalent in operations that process physical items is transportation costs. For example, if a manufacturer supplies the whole of its European market from one major plant in Denmark, all supplies may have to be brought in from several countries to the single plant and all products shipped from there throughout Europe.

> **OPERATIONS PRINCIPLE**
> ***Diseconomies of scale increase operating costs above a certain level of capacity, resulting in a minimum cost-level of capacity.***

Being small may have advantages

Although large-scale-capacity operations will usually have a cost advantage over smaller units, there are also potentially significant advantages that can be exploited by small-scale operations. One significant research study showed that small-scale operations can provide significant advantages in the following four areas:[7]

1. Locating near to 'hot spots' that can tap into local knowledge networks. Often, larger companies centralise their research and development efforts, so losing touch with where innovative ideas are generated.
2. Responding rapidly to regional customer needs and trends by basing more and smaller units of capacity close to local markets.
3. Taking advantage of the potential for human resource development, by allowing staff a greater degree of local autonomy. Larger-scale operations often have longer career paths with fewer opportunities for 'taking charge'.
4. Exploring radically new technologies by acting in the same way as a smaller more entrepreneurial rival. Larger, more centralised development activities are often more bureaucratic than smaller-scale agile centres of development.

The timing of capacity change

Changing the capacity of any operation in a supply network is not just a matter of deciding on its optimum capacity. The operation also needs to decide when to bring 'on-stream' new capacity. For example, Figure 4.9 shows the forecast demand for a manufacturer's new product. In deciding *when* new capacity is to be introduced the company can mix three strategies:

1. Capacity is introduced to generally lead demand, timing the introduction of capacity in such a way that there is always sufficient capacity to meet forecast demand.
2. Capacity is introduced to generally lag demand, timing the introduction of capacity so that demand is always equal to or greater than capacity.
3. Capacity is introduced to sometimes lead and sometimes lag demand, but inventory built up during the 'lead' times is used to help meet demand during the 'lag' times. This is called 'smoothing with inventory'.

Each strategy has its own advantages and disadvantages. These are shown in Table 4.2. The actual approach taken by any company will depend on how it views these advantages and disadvantages. For example, if the company's access to funds for capital expenditure is limited, it is likely to find the delayed capital expenditure requirement of the capacity-lagging strategy relatively attractive. Of course, the third strategy, 'smoothing with inventory', is only appropriate for operations that produce products that can be stored. Customer-processing operations such as a hotel cannot satisfy demand in one year by using rooms that were vacant the previous year.

> **OPERATIONS PRINCIPLE**
> ***Capacity-leading strategies increase opportunities to meet demand. Capacity-lagging strategies increase capacity utilisation.***

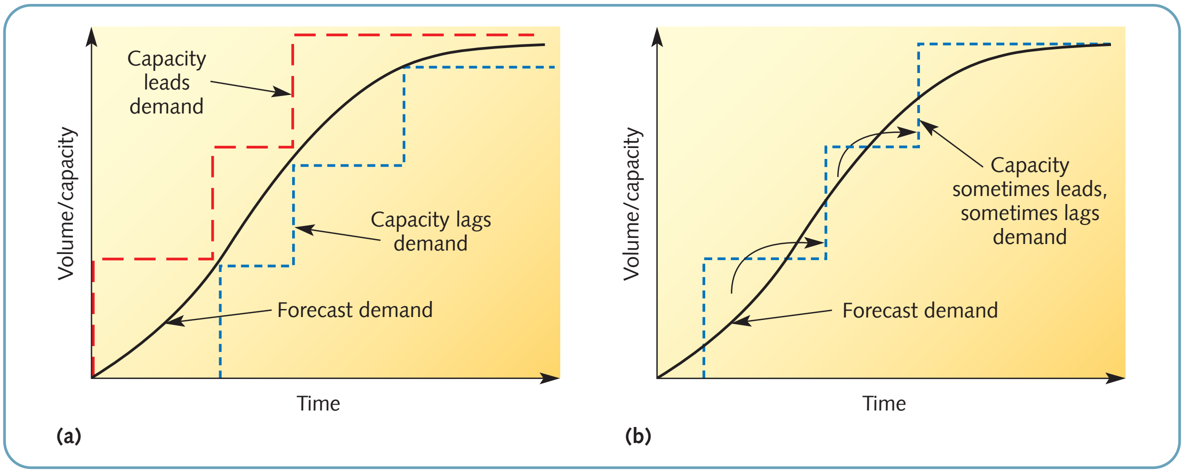


Figure 4.9 Capacity-leading and capacity-lagging strategies; (a) smoothing with inventories (b) means using the excess capacity in one period to produce inventory that supplies the under-capacity period

Source: From Slack, N. *Operations Management*, *8e*, © 2016 Pearson Education Limited, UK.

Table 4.2 The arguments for and against pure leading, pure lagging and smoothing with inventory strategies of capacity timing

Advantages	*Disadvantages*
Capacity-leading strategies	
Always sufficient capacity to meet demand, therefore revenue is maximised and customers are satisfied	Utilisation of the plants is always relatively low, therefore costs will be high
Most of the time there is a 'capacity cushion' that can absorb extra demand if forecasts are pessimistic	Risks of even greater (or even permanent) over-capacity if demand does not reach forecast levels
Any critical start-up problems with new operations are less likely to affect supply	Capital spending on capacity will be early
Capacity-lagging strategies	
Always sufficient demand to keep the operation working at full capacity, therefore unit costs are minimised	Insufficient capacity to meet demand fully, therefore reduced revenue and dissatisfied customers
Over-capacity problems are minimised if forecasts prove optimistic	No ability to exploit short-term increases in demand
Capital spending on the operation is delayed	Under-supply position even worse if there are start-up problems with the new operations
Smoothing with inventory strategies	
All demand is satisfied, therefore customers are satisfied and revenue is maximised	The cost of inventories in terms of working capital requirements can be high; this is especially serious at a time when the company requires funds for its capital expansion
Utilisation of capacity is high and therefore costs are low	Risks of product deterioration and obsolescence
Very short-term surges in demand can be met from inventories	

Break-even analysis of capacity expansion

An alternative view of capacity expansion can be gained by examining the cost implications of adding increments of capacity on a break-even basis. Figure 4.10 shows how increasing capacity can move an operation from profitability to loss. Each additional unit of capacity results in a *fixed-cost break* that is a further lump of expenditure that will have to be incurred before any further activity can be undertaken in the operation. The operation is unlikely to be profitable at very low levels of output. Eventually, assuming that prices are greater than marginal costs, revenue will exceed total costs. However, the level of profitability at the point where the output level is equal to the capacity of the operation may not be sufficient to absorb all the extra fixed costs of a further increment in capacity. This could make the operation unprofitable in some stages of its expansion.

> OPERATIONS PRINCIPLE
> ***Using inventories to overcome demand–capacity imbalance tends to increase working capital requirements.***

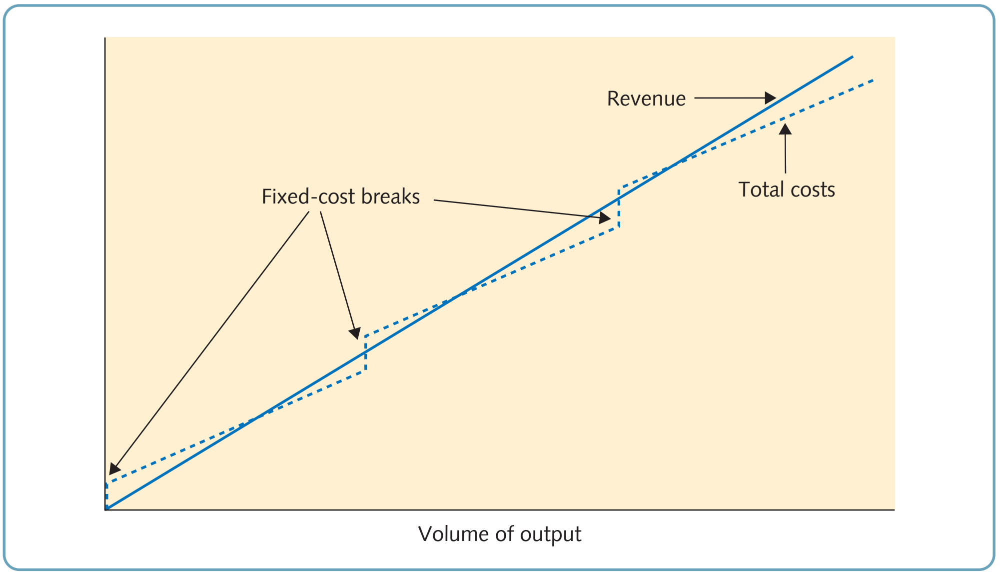


Figure 4.10 Incurring fixed costs repeatedly can raise total costs above revenue over some ranges of output

Worked example

De Vere Graphics

De Vere Graphics is investing in a new machine that enables it to make high-quality prints for its clients. Demand for these prints is forecast to be around 100,000 units in year 1 and 220,000 units in year 2. The maximum capacity of each machine the company will buy to process these prints is 100,000 units per year. They have a fixed cost of €200,000 per year and a variable cost of processing of €1 per unit. The company believes it will be able to charge €4 per unit for producing the prints. What profit is it likely to make in the first and second years?

Year 1 demand = 100,000 units; therefore the company will need one machine

Cost of producing prints = fixed cost for one machine + variable cost × 100,000
= €200,000 + (€1 × 100,000)
= € 300,000

Revenue = demand × price
= 100,000 × €4
= €400,000

Therefore profit = €400,000 − €300,000
= €100,000

Year 2 demand = 220,000; therefore the company will need three machines

Cost of manufacturing = fixed cost for three machines + variable cost × 220,000
= (3 × €200,000) + (€1 × 220,000)
= €820,000

Revenue = demand × price
= 220,000 × €4
= €880,000

Therefore profit = €880,000 − €820,000
= €60,000

Note: the profit in the second year will be lower because of the extra fixed costs associated with the investment in the two extra machines.

4.6 Diagnostic question: *Where should operations be located?*

The location of each operation in a supply network is a key element in defining its structure, and also will have an impact on how the network operates in practice. Poor location of any operation in a supply network can have a significant impact, not just on its profits, but also those of others in the network. For example, siting a data centre where potential staff with appropriate skills will not live will affect its performance, and the service it gives its customers. Location decisions will usually have an effect on an operation's costs as well as its ability to serve its customers (and therefore its revenues). In addition, location decisions, once taken, are difficult to undo. The costs of moving an operation can be hugely expensive and the risks of inconveniencing customers very high. No operation wants to move very often.

Why relocate?

Not all operations can logically justify their location. Some are where they are for historical reasons. Yet, even the operations that are 'there because they're there' are implicitly deciding not to move. Presumably, their assumption is that the cost and disruption involved in changing location would outweigh any potential benefits of a new location. When operations do move, it is usually for one or both of two reasons – changes in demand or changes in supply.

Changes in demand

If customer demand shifts it may prompt a change in location. For example, as garment manufacturers moved to Asia, suppliers of zips, threads and so on started to follow them. Changes in the volume of demand can also prompt relocation. To meet higher demand,

Case example

Aerospace in Singapore

It is not immediously obvious why Singapore has been so successful at attracting a significant proportion of Asia's aerospace business. Unlike most states in the region, it has practically no internal air transport, but it does have one of the most respected airlines in the world in Singapore International Airlines (SIA). Yet it has attracted such global aerospace firms as Airbus, Rolls-Royce, Pratt and Whitney, Thales, Bombardier and many more. There are a number of reasons for this.[8] First, the location had access to the skills and infrastructure to support technically complex manufacturing. It was ranked number one globally in the 2019 World Economic Forum Global Competitiveness Report. Second, the country was trusted to provide an ethical framework for business. Again, it was ranked first in Asia for intellectual property rights protection in the 2019 World Economic Forum Global Competiveness Report.[9] Third, Asia is where the demand is. The world's fastest-growing airlines are in China, Singapore, Indonesia, India and in the Gulf. Fourth, the Singapore government offered significant help for companies wanting to invest in the sector, including generous tax incentives. Yet, although important, these incentives were not as beneficial as the 'soft' factors that make Singapore so attractive. In particular, the City State's universities and colleges, which produce the skilled scientists, engineers and staff who are vital to producing products that cannot be allowed to fail. The talent pipeline that the sector needed was seen as excellent and sustainable. Aerospace engineering courses are popular at Singapore's institutes of higher learning, and aerospace companies partner schools to develop courses and offer attachments and on-the-job training. Nor is government encouragement confined to large international aerospace firms. Many small- and medium-sized enterprises in the sector have built technical capabilities, particularly in areas such as aircraft and engine parts manufacturing, providing aerospace component maintenance, repair and overhaul (MRO) services.

an operation could expand its existing site, or choose a larger site in another location, or keep its existing location and find a second location for an additional operation; the last two options will involve a location decision. High-visibility operations may not have the choice of expanding on the same site to meet rising demand. A dry-cleaning service may attract only marginally more business by expanding an existing site because it offers a local, and therefore convenient, service. Finding a new location for an additional operation is probably its only option for expansion.

Changes in supply

The other stimulus for relocation is changes in the cost, or availability, of the supply of inputs to the operation. For example, a mining or oil company will need to relocate as the minerals it is extracting become depleted. The reason why so many software companies located in India was the availability of talented, well-educated but relatively cheap staff.

Evaluating potential changes in location

Evaluating possible locations is almost always a complex task because the number of location options, the criteria against which they could be evaluated and the comparative rarity of a single location that clearly dominates all others, make the decision strategically sensitive. Furthermore, the decision often involves high levels of uncertainty. Neither the relocation activity itself, nor the operating characteristics of the new site could be as assumed when the decision was originally made. Because of this, it is useful to be systematic in terms of (a) identifying alternative options, and (b) evaluating each option against a set of rational criteria.

OPERATIONS PRINCIPLE
An operation should only change its location if the benefits of moving outweigh the costs of operating in the new location plus the cost of the move itself.

Case example

IKEA closes a seven-storey store[10]

Sometimes it is the physical size, positioning and shape of an operation's site that is just as (or more) important than its wider geographic location. Even IKEA, a company with an outstanding reputation for operations management, learned this lesson when it closed its Coventry store in the UK, more than ten years after it opened. Although IKEA was experimenting with different types of store than its traditional large-area stores (see the end-of-chapter case study in Chapter 2), the reason for closure was not that big stores were losing their appeal, as the company announced a large new store in Hammersmith, West London almost simultaneously. Part of the issue with the Coventry store was its location in the centre of the city. Customers had been showing a preference for shopping in retail parks and online, which had resulted in reducing customer numbers. The design of the store had also proved problematic: its city-centre location had resulted in a relatively cramped site, which meant that the store had been built over seven levels in order to reach the standard IKEA shop size of 2,700 square metres. This had had a significant effect on the operating costs of the store, as well as the shopping experience for customers. The company said that it had not been able either to make the store more cost-efficient or to repurpose it. Nor was it a recent problem. Although thousands of customers had queued up to shop in the store when it opened, with extra police called in to control the crowds, IKEA said that the shop had actually lost money consistently.

Identify alternative location options

The first relocation option to consider is not to relocate. Sometimes relocation is inevitable, but often staying put is a viable option. It is worth evaluating the 'do nothing' option, if only to provide a 'base case' against which to compare other options. But in addition to the 'do nothing' option there should be a number of alternative location options. The implication of the globalisation of the location decision has been to increase both the number of options and the degree of uncertainty in their relative merits. The sheer number of possibilities makes the location decision impossible to 'optimise'. Rather, the process of identifying location options usually involves selecting a limited number of sites that represent different attributes. For example, a distribution centre, while always needing to be close to transport links, could be located in any of several regions and could either be close to population centres, or in a more rural location. The options may be chosen to reflect a range of both these factors. However, this assumes that the 'supply' of location options is relatively large, which is not always the case. In many retail location decisions, there are a limited number of high-street locations that become available at any point in time. Often, a retailer will wait until a feasible location becomes available and then decide whether to either take up that option or wait and take the chance that a better location becomes available soon. In effect, the location decision here is a sequence of 'take or wait' decisions.

Set location evaluation criteria

Although the criteria against which alternative locations can be evaluated will depend on specific circumstances, the following five broad categories are typical:

1. ***Capital requirements*** – the capital or leasing cost of a site is usually a significant factor. This will probably be a function of the location of the site and its characteristics, such as the shape of the site, site access, the availability of utilities, and so on.
2. ***Market factors*** – location can affect how the market, either in general, or as individual customers, perceives an operation. Locating a general hospital in the middle of the countryside may have many advantages for its staff, but it clearly would be very inconvenient for its customers. Likewise, restaurants, stores, banks, petrol filling stations and many other

high-visibility operations must all evaluate how alternative locations will determine their image and the level of service they can give. The same arguments apply to labour markets. Location may affect the attractiveness of the operation in terms of staff recruitment and retention. But not all locations necessarily have appropriate skills available immediately. Staff at a remote call centre in the western islands of Scotland, who were used to a calm and tranquil life, became stunned by the aggressive nature of many callers to the call centre. They had to be given assertiveness training by the call-centre management.

3. ***Cost factors*** – two major categories of cost are affected by location. The first is the cost of producing products or services. For example, labour costs can vary between different areas in any country, but are likely to be a far more significant factor when international comparisons are made; they can exert a major influence on the location decision, especially in some industries such as clothing, for example, where labour costs as a proportion of total costs are relatively high. Other cost factors, known as community factors, derive from the social, political and economic environment of the site. These include such factors as local tax rates, capital movement restrictions, government financial assistance, political stability, local attitudes to 'inward investment', language, local amenities (schools, theatres, shops, etc.), the availability of support services, the history of labour relations and behaviour, environmental restrictions and planning procedures. The second category of costs relates to both the cost of transporting inputs from their source to the location of the operation and the cost of transporting products and services from the location to customers. Whereas almost all operations are concerned to some extent with the former, not all operations are concerned with the latter, either because customers come to them (for example, hotels), or because their services can be 'transported' at virtually no cost (for example, technology helpdesks). For supply networks that process physical items, however, transportation costs can be very significant.

4. ***Future flexibility*** – because operations rarely change their location, any new location must be capable of being acceptable, not only under current circumstances, but also under possible future circumstances. The problem is that no one knows exactly what the future holds. Nevertheless, especially in uncertain environments, any evaluation of alternative locations should include some kind of scenario planning that considers the robustness of each in coping with a range of possible futures.

5. ***Risk factors*** – closely related to the concept of future flexibility, is the idea of evaluating the risk factors associated with possible locations. Long-term risks could include damaging changes in input factors such as exchange rates or labour costs, but can also include more fundamental security risks to staff or property.

Critical commentary

- Probably the most controversial issue in supply network design is that of outsourcing. In many instances, there has been fierce opposition to companies outsourcing some of their processes. Trade unions often point out that the only reason that outsourcing companies can do the job at lower cost is that they either reduce salaries, reduce working conditions, or both. Furthermore, they say, flexibility is only achieved by reducing job security. Employees who were once part of a large and secure corporation could find themselves as far less-secure employees of a less-benevolent employer with a philosophy of permanent cost cutting. Even some proponents of outsourcing are quick to point out the problems. There can be significant obstacles, including understandable resistance from staff who find themselves 'outsourced'. Some companies have also been guilty of 'outsourcing a problem'. In other words, having failed to manage a process well themselves, they ship it out rather than face up to why the process was problematic in the first place. There is also evidence that, although long-term costs can be brought down when a process is outsourced, there may be an initial period when costs rise as both sides learn how to manage the new arrangement.

- The idea of widening the discussion of supply networks to include the 'business ecosystem' concept, described earlier, is also not without its critics. Some see it as simply another management 'buzzword', indistinguishable from the longer-established idea of the supply network. Other critics, who believe that the ecosystem metaphor is just a way for businesses to appear 'green', have criticised the use of the term 'business ecosystem' by commentators and firms. They claim that the metaphor is used to suggest that the commercial relationships, on which almost all supply networks are based, have developed and are run using 'natural' values and therefore should be left to operate free from societal or government interference.

SUMMARY CHECKLIST

- ☐ Is the operation fully aware of all its first- and second-tier suppliers' and customers' capabilities and requirements?
- ☐ Are the capabilities of suppliers and requirements of customers understood in terms of all aspects of operations performance?
- ☐ Does the operation have a view on how it would like to see its supply network develop over time, both in terms of scope and structure?
- ☐ Have the benefits of reducing the number of individual suppliers been explored?
- ☐ Are any parts of the supply network likely to become disintermediated, and have the implications of this been considered?
- ☐ Does the operation have an approach to how it treats others in the supply network who might be both complementors and competitors?
- ☐ Is the vertical integration/outsourcing issue always under review for possible benefits?
- ☐ Is outsourcing (or bringing back in-house) evaluated in terms of all the operation's performance objectives?
- ☐ Is there a rational set of criteria used for deciding whether (or not) to outsource?
- ☐ Is the optimum economy of scale for the different types of operation within the business periodically assessed?
- ☐ Are the various strategies for timing changes in capacity always evaluated in terms of their advantages and disadvantages?
- ☐ Are the fixed-cost breaks of capacity increase understood, and are they taken into account when increasing or decreasing capacity?
- ☐ Is the relocation decision ever considered?
- ☐ Have factors such as changes in demand or supply that may prompt relocation been considered?
- ☐ If considering relocation, are alternative locations always evaluated against each other and against a 'do nothing' option?
- ☐ Are sufficient location options being considered?
- ☐ Do location evaluation criteria include capital, market, cost, flexibility and risk factors?

Case study

Aarens Electronic

Just outside Rotterdam in the Netherlands, Francine Jansen, the Chief Operating Officer of Aarens Electronic (AE) was justifiably proud of what she described as '*the most advanced machine of its type in the world, which will enable us to achieve new standards of excellence for our products requiring absolute cleanliness and precision*' and '*a quantum leap in harnessing economies of scale, new technology to provide **the** most advanced operation for years to come*'. The Rotterdam operation was joining AE's two existing operations in the Netherlands. It offered precision custom coating and laminating services to a wide range of customers, among the most important being Phanchem, to whom it supplied dry photoresist imaging films, a critical step in the manufacturing of microchips. Phanchem then processed the film further and sold it direct to microchip manufacturers

The Rotterdam operation

The decision to build the Rotterdam operation had been taken because the company believed that a new low-cost operation using 'ultra-clean' controlled environment technology could secure a very large part of Phanchem's future business – perhaps even an exclusive agreement to supply 100 per cent of its needs. When planning the new operation, three options were presented to AE's Executive Committee:

1. Expand an existing site by building a new machine within existing site boundaries. This would provide around 12–13 million square metres (MSM) per year of additional capacity and require around €19 million in capital expenditure.
2. Build a new facility alongside the existing plant. This new facility could accommodate additional capacity of around 15 MSM per year but, unlike option A, would also allow for future expansion. Initially, this would require around €22 million of capital.
3. Set up a totally new site with a much larger increment of capacity (probably around 25 MSM per year). This option would be more expensive – at least €30 million.

Francine Jansen and her team initially favoured option 2 but in discussion with the AE Executive Committee, opinion shifted towards the more radical option 3: '*It may have been the highest-risk option but it held considerable potential and it fitted with the AE Group philosophy of getting into high-tech specialised areas of business. So we went for it.*' (Francine Jansen) The option of a very large, ultra-clean, state-of-the-art facility also had a further advantage: it could change the economics of the photoresist imaging industry. In fact, global demand and existing capacity did not immediately justify investing in such a large increase in capacity. There was probably some over-capacity in the industry. But a large-capacity, ultra-clean operation could provide a level of quality at such low costs that, if there were over-capacity in the industry, it would not be AE's capacity that would be lying idle.

Designing the new operation

During discussions on the design of the new operation, it became clear that there was one issue that was underlying all the team's discussions – how flexible should the process be? Should the team assume that they were designing an operation that would be dedicated exclusively to the manufacture of photoresist imaging film, and ruthlessly cut out any technological options that would enable it to manufacture other products, or should they design a more general-purpose operation that was suitable for photoresist imaging film, but could also make other products? It proved a difficult decision. The advantages of the more flexible option were obvious: '*At least it would mean that there was no chance of me being stuck with an operation and no market for it to serve in a couple of years' time.*' (Francine Jansen) But the advantages of a totally dedicated operation were less obvious, although there was a general agreement that both costs and quality could be superior in an operation dedicated to one product.

Eventually, the team decided to focus on a relatively non-flexible focused and dedicated large machine. '*You can't imagine the agonies we went through when we decided not to make this a flexible machine. Many of us were not comfortable with saying "this is going to be a photoresist machine exclusively, and if the market goes away we're in real trouble". We had a lot of debate about that. Eventually, we more or less reached a consensus for focus but it was certainly one of the toughest decisions we ever made.*' (Francine Jansen) The capital cost savings of a focused facility and operating costs savings of up to 25 per cent were powerful arguments, as was the philosophy of total process dedication: '*The key word for us was **focus**. We wanted to be quite clear about what was needed to satisfy our customer in making this single type of product. As well as providing significant cost savings to us, it made it a lot easier to identify the root causes of any problems because we would not have to worry about how it might affect other products. It's all very clear. When the line was down we would not be generating revenue! It would also force us to understand our own performance. At our other operations, if a line goes down, the people can be shifted to other responsibilities. We don't have other responsibilities here – we're either making it or we're not.*' (Francine Jansen)

When the Rotterdam operation started producing, the team had tweaked the design to bring the capacity at start-up to 32 MSM per year. And notwithstanding some initial teething troubles it was, from the start, a technical and commercial success. Within six months a contract was signed with Phanchem to supply 100 per cent of Phanchem's needs for the next 10 years. Phanchem's decision was based on the combination of manufacturing and business focus that the Rotterdam team had achieved, a point stressed by Francine Jansen: '*Co-locating all necessary departments on the Rotterdam site was seen as particularly important. All the technical functions and the marketing and business functions are now on site.*'

Developing the supply relationship

At the time of the start-up, product produced in Rotterdam was shipped to Phanchem's facility near Frankfurt, Germany, almost 500 km away. This distance caused a number of problems, including some damage in transit and delays in delivery. However, the relationship between AE and Phanchem remained sound, helped by the two companies' cooperation during the Rotterdam start-up. '*We had worked closely with them during the design and construction of the new Rotterdam facility. More to the point, they saw that they would certainly achieve cost savings from the plant, with the promise of more savings to come as the plant moved down the learning curve.*' (Francine Jansen) The closeness of the relationship between the two companies was a result of their staff working together. AE engineers were impressed by their customer's willingness to help out while they worked on overcoming the start-up problems. Similarly AE had helped Phanchem when it needed extra supplies at short notice. As Francine Jansen said, '*partly because we worked together on various problems the relationship has grown stronger and stronger*'.

In particular, the idea of a physically closer relationship between AE and Phanchem was explored. '*During the negotiations with Phanchem for our 100 per cent contract there had been some talk about co-location, but I don't think anyone took it particularly seriously. Nevertheless, there was general agreement that it would be a good thing to do. After all, our success as Phanchem's sole supplier of coated photoresist was tied in to their success as a player in the global market; what was good for Phanchem was good for AE.*' (Francine Jansen) Several options were discussed within and between the two companies. Phanchem had, in effect, to choose between four options:

1. Stay where they were near Frankfurt.
2. Relocate to the Netherlands (which would give easier access to port facilities) but not too close to AE (an appropriate site was available 30 km from Rotterdam).
3. Locate to a currently vacant adjacent site across the road from AE's Rotterdam plant.
4. Co-locate within an extension that could be specially built onto the AE plant at Rotterdam.

Evaluating the co-location options

Relatively early in the discussions between the two companies, the option of 'doing nothing' by staying in Frankfurt was discounted. Phanchem wanted to sell its valuable site near Frankfurt. The advantages of some kind of move were significant. The option of Phanchem moving to a site 30 km from Rotterdam was considered but rejected because it had no advantages over locating even closer to the Rotterdam plant. Phanchem also strongly considered building and operating a facility across the road from the Rotterdam plant. But eventually the option of locating in a building attached to AE's Rotterdam operation became the preferred option. Co-location would have a significant impact on Phanchem's competitiveness by reducing its operating costs, enabling it to gain market share by offering quality film at attractive prices, thus increasing volume for AE. The managers at the Rotterdam plant also looked forward to an even closer operational relationship with the customer: '*Initially, there was some resistance in the team to having a customer on the same site as ourselves. No one in AE had ever done it before. The step from imagining our customer across the road to imagining them on the same site took some thinking about. It was a matter of getting used to the idea, taking one step at a time.*' (Francine Jansen)

The customer becomes a paying guest

However, when Francine and the Rotterdam managers presented their proposal for extending the plant to the AE board the proposal was not well received. '*Leasing factory space to our customer seemed a long way from our core business. As one executive committee member said, "we are manufacturers; we aren't in the real estate business". But we felt that it would be beneficial for both companies.*' (Francine Jansen) And even when the proposal was eventually accepted, there was still concern over sharing a facility. In fact the Executive Committee insisted that the door between the two companies' areas should be capable of being locked from both sides. Yet the construction and commissioning of the new facility for Phanchem was also a model of cooperation. Now, all visitors to the plant are shown the door that had to be 'capable of being locked from both sides' and asked how many times they think it has been locked. The answer, of course, is 'never'.

Questions

1. What were the key structure and scope decisions taken by Aarens Electronic?
2. What were the risks involved in adopting a process design that was 'totally dedicated' to the one customer's needs?
3. What were the advantages and disadvantages of each location option open to Phanchem, and why do you think it eventually chose to co-locate with AE?

Applying the principles

Some of these exercises can be answered by reading the chapter. Others will require some general knowledge of business activity and some might require an element of investigation. **Model answers for the first two questions are available on the student companion website.** *For model answers to the other questions in this section, please ask your tutor/lecturer.*

1. Consider the music business as a supply network. How did music downloads and streaming affect artists' sales? What implications did electronic music transmission have for record shops?

2. A data centre is 'a facility composed of networked computers and storage that businesses or other organisations use to organise, process, store and disseminate large amounts of data. A business typically relies heavily upon the applications, services and data contained within a data centre, making it a focal point and critical asset for everyday operations.'[11] These facilities can contain network equipment, servers, data storage and back-up facilities, software applications for large companies, and more. Very few businesses (or people) do not rely on them. And determining their location is a crucial decision for those operations running them. In fact, such businesses usually have a set method for choosing data centre location. Visit the websites of the types of businesses that run data centres (such as Intel, Cisco or SAP) and devise a set of criteria that could be used to evaluate potential sites.

3. A company that produces concrete paving slabs is introducing a new range of 'textured' non-slip products. To do this, it must invest in a new machine. Demand is forecast to be around 10,000 units per month for the first year and approximately 24,000 units per month after that. The machines that produce these products have a capacity of 10,000 units per month. They have a fixed cost of €20,000 per month and a variable cost of processing of €1 per unit. The company has forecast that they will be able to charge €4 per unit. It has been suggested that they would make higher profits if sales were restricted to 20,000 units per month in the second year. Is this true?

4. The Fast and Efficient (FAE) Transport Group is reviewing its fleet maintenance operations: '*Our lease on our current maintenance and repair facilities site will expire in a year, and we need to decide how to operate in the future. Currently, we have the one site with five repair bays. This can cope with our fleet of 40 trucks. But demand is growing; and within two or three years we hope to be operating around 55–60 trucks, so we will have to choose a site (or sites) that allows for this increase. And that leads me to the next issue – should we stick to operating one central site, or should we plan to have two sites, one for the north and one for the south of our region?*'

   As far as FAE's operations managers could forecast, the costs of having one or two sites would be as follows:

   One site – fixed cost of establishing the site = € 300,000; variable cost of servicing trucks = € 300,000 per truck per year.

   Two sites – fixed cost of establishing the sites (for both) = € 500,000; variable cost of servicing trucks = € 10,000 per truck per year (they will be out of action for less time because sites would be closer).

   At what level of demand (in terms of the number of trucks operated by the company) will the two-site proposal be cheaper?

5. Globalisation is very much a 'mixed blessing'. There is little doubt that it has lifted millions out of poverty, but it has also led to the destruction of traditional cultures in developing countries and many jobs in the developed world. Draw up lists of what you see as the advantages and disadvantages of globalisation.

Notes on chapter

1 Christopher, M. (2011) 'Logistics and supply chain management', *Financial Times,* London.

2 Fitzgerald, F. Scott (1941) *The Last Tycoon*, Penguin Classics, 2001.

3 The information on which this example is based is taken from: the Compass website, www.compass-group.com/en [accessed 14 September 2020]; Fildes, N. (2017) 'Vodafone to bring 2,100 call-centre jobs back to UK', *Financial Times*, 13 March; Flinders, K. (2017) 'Vodafone brings offshore contact centre work to UK', *Computer Weekly*, 13 March.

4 For further details of the origin of this concept see: Moore, J.F. (1993) 'Predators and prey: A new ecology of competition', *Harvard Business Review,* May–June.

5 Moore, J.F. (1993) *op. cit.*

6 The information on which this example is based is taken from: Margolis, J. (2019) 'Amazon's 4-Star stores could be the future of bricks-and-mortar retail', *Financial Times*, 24 October; Economist (2016) 'Shops to showrooms: Why some firms are opening shops with no stock', *The Economist* print edition, 12 March.

7 Bacon, G., Machan, I. and Dnyse, J. (2008) 'Offshore challenges, manufacturing', *Institute of Electrical Engineers,* January.

8 There are many papers and articles outlining Singapore's aerospace strategy. For example, see Singapore EDB (2020) 'Singapore: Asia's aerospace hub', EDB Aerospace Industry brochure; Choo Yun Ting (2020) 'Support for SMEs to help aerospace industry soar', *The Straits Times*, 17 February; Raghuvanshi, G. (2013) 'Rolls-Royce Pushes Focus on Singapore', *Wall Street Journal,* 15 September.

9 Schwab, K. (2019) 'The Global Competitiveness Report 2019', Insight Report, World Economic Forum, available at http://www3.weforum.org/docs/WEF_TheGlobalCompetitivenessReport2019.pdf [accessed 14 September 2020].

10 The information on which this example is based is taken from: Armstrong, A. (2020) 'Ikea blames bad design for megastore closure', *The Times*, 5 February; Partridge, J. (2020) 'Ikea to close first big UK store, putting 350 jobs at risk', *The Guardian*, 4 February.

11 Definition from TechTarget.com, see searchdatacenter.techtarget.com [accessed 29 September 2020].

Taking it further

Chopra, S. and Meindl, P. *(2014) Supply Chain Management: Strategy, planning and operations, 6th edition,* Prentice Hall. *A good textbook that covers both strategic and operations issues*.

Corbett, M.F. *(2010) The Outsourcing Revolution: Why it makes sense and how to do it right,* Kaplan. *Not an academic book on outsourcing.*

Cullen S.K., Lacity, M. and Willcocks L.P. *(2014) Outsourcing: All You Need To Know,* White Plume Publishing. *A practical guide to outsourcing, with lots of examples across sectors and countries, from an experienced team of authors.*

Gunasekaran, A., Lai, K.-H. and Edwin Cheng, T.C. *(2008) 'Responsive supply chain: A competitive strategy in a networked economy', Omega, 36 (4), pp. 549–564.* *An academic paper, but explains it well.*

Kim, Y., Chen, Y.-S. and Linderman, K. *(2015) 'Supply network disruption and resilience: A network structural perspective', Journal of Operations Management, 33–34 (1), pp. 43–59.* *A more in-depth article examining the effect of different structural networks on the likelihood of disruptions.*

Moore, J.F. ***(2013) Shared Purpose: A thousand business ecosystems, a connected community, and the future, CreateSpace Independent Publishing Platform.*** *A short but sweet progress report on a study of ARM Holdings and its 1,000-plus community of partners from the founder of the business ecosystems idea.*

Oshri, I., Kotlarsky, J. and Willcocks, L.P. ***(2015) The Handbook of Global Outsourcing and Offshoring, 3rd edition, Palgrave Macmillan.*** *A clear and applied guide to the key steps in outsourcing and offshoring decisions.*

Singh Srai, J. and Gregory, M. ***(2008) 'A supply network configuration perspective on international supply chain development', International Journal of Operations & Production Management, 28 (5), pp. 386–411.*** *Another paper for an academic audience, but nevertheless worth reading.*

Steger, M.B. ***(2017) Globalization: A very short introduction, Oxford University Press.*** *Not just a business book, this also covers political, cultural, ideological and environmental issues.*

Supplement to Chapter 4

Forecasting – Does the operation have an idea of future demand?

Introduction

Some forecasts are accurate. We know exactly what time the sun will rise at any given place on Earth tomorrow, or one day next month, or even next year. Forecasting in a business context, however, is far more difficult and therefore prone to error. We do not know precisely how many orders we will receive or how many customers will walk through the door tomorrow, next month or next year. But such forecasts are necessary to help managers make decisions about resourcing the organisation for the future.

Forecasting – knowing the options

Simply knowing that demand for your goods or services is rising or falling is not enough in itself; knowing the rate of change is likely to be vital to business planning. A firm of lawyers may have to decide the point at which, in its growing business, it will have to take on another partner. Hiring a new partner could take months so it needs to be able to forecast when it expects to reach that point and then when it needs to start its recruitment drive. The same applies to a plant manager who will need to purchase new plant to deal with rising demand. She may not want to commit to buying an expensive piece of machinery until absolutely necessary, but in enough time to order the machine and have it built, delivered, installed and tested. The same is so for governments, whether planning new airports or runway capacity, or deciding where and how many primary schools to build.

The first question is to know how far you need to look ahead and this will depend on the options and decisions available to you. Take the example of a local government, where the number of primary-age children (4–11 years) is increasing in some areas and declining in other areas within its boundaries. It is legally obliged to provide school places for all such children. Government officials will have a number of options open to them and they may each have different lead times associated with them. One key step in forecasting is to know the possible options and the lead times required to bring them about (see Table 4.3).

Individual schools can hire (or lay-off) short-term (supply) teachers from a pool, not only to cover for absent teachers, but also to provide short-term capacity while teachers are hired

Table 4.3 Options available and lead time required for dealing with changes in numbers of school children

Options available	*Lead time required*
Hire short-term teachers	Hours
Build temporary classrooms	↓
Hire permanent staff	↓
Amend school catchment areas	↓
Build new classrooms	↓
Build new schools	Years

to deal with increases in demand. Acquiring (or dismissing) such temporary cover may only require a few hours' notice. (This is often referred to as 'short-term capacity management'.)

A shortage of accommodation may be fixed in the short to medium term by hiring or buying temporary classrooms. It may take only a couple of weeks to hire such a building and equip it ready for use.

Hiring new permanent staff or laying-off existing staff is another option, but this kind of action may take months to complete ('medium-term capacity management').

It may be possible to amend catchment areas between schools to try to balance an increasing population in one area against a declining population in another. However, such changes often require lengthy consultation processes.

In the longer term, new classrooms or even new schools may have to be built. The planning, consultation, approval, commissioning, tendering, building and equipping process may take one to five years, depending on the scale of the new build ('long-term capacity management').

Knowing the range of options, managers can then decide the timescale for their forecasts; indeed, several forecasts might be needed for the short term, medium term and long term.

In essence, forecasting is simple

To know how many children may turn up in a local school tomorrow, you can use the number that turned up today. And in the long term, in order to forecast how many primary-aged children will turn up at a school in five years' time, one need simply look at the birth statistics for the current year for the school's catchment area (see Figure 4.11).

However, such simple extrapolation techniques are prone to error and indeed such approaches have resulted in some local governments committing themselves to building schools that five or six years later, when complete, had few children, while other schools were bursting at the seams with temporary classrooms and temporary teachers, often resulting in falling morale and declining educational standards. The contextual variables (see Figure 4.12) that

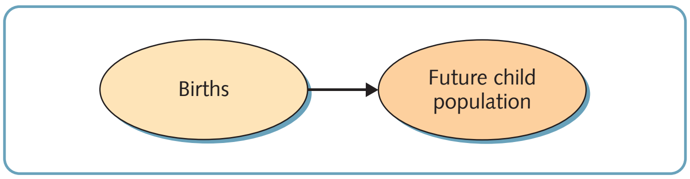


Figure 4.11 Simple prediction of future child population

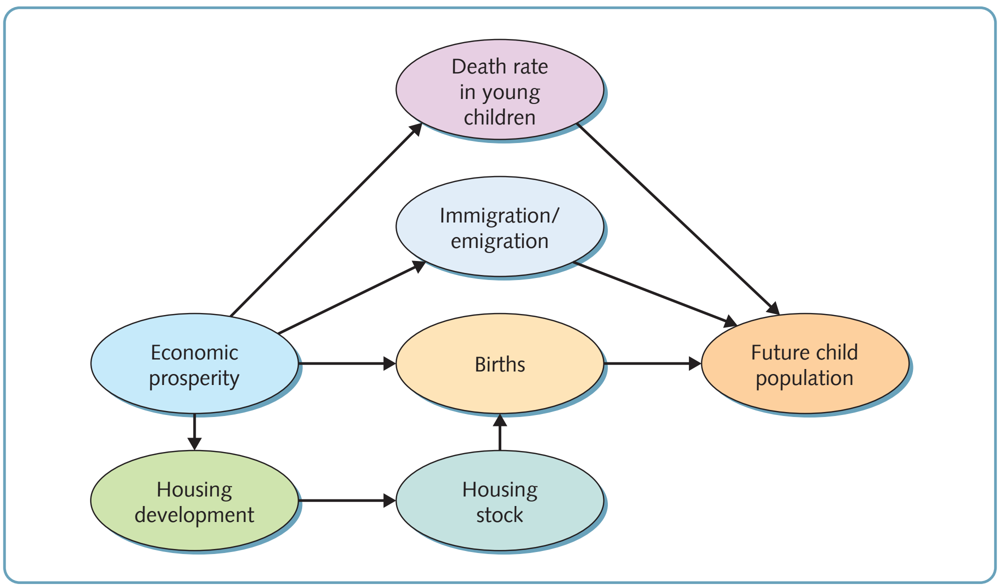


Figure 4.12 Some of the key causal variables in predicting child population

will have a potentially significant impact on, for example, the school population five years hence include the following:

- One minor factor in developed countries, though a major factor in developing countries, might be the death rate in children between birth and five years of age. This may be dependent upon location, with a slightly higher mortality rate in the poorer areas compared to the more affluent areas.
- Another more significant factor is immigration and emigration, as people move into or out of the local area. This will be affected by housing stock and housing developments, the ebb and flow of jobs in the area and changing economic prosperity in the area.
- A key factor that has an impact on the birth rate in an area is the amount and type of the housing stock. City-centre tenement buildings tend to have a higher proportion of children per dwelling, for example, than suburban semi-detached houses. So not only will existing housing stock have an impact on the child population, but so also will the types of housing developments proposed, planned or under construction.

Approaches to forecasting

There are two main approaches to forecasting. Managers sometimes use qualitative methods based on opinions, past experience and even best guesses. There is also a range of qualitative forecasting techniques available to help managers evaluate trends and causal relationships, and make predictions about the future. Secondly, quantitative forecasting techniques can be used to model data. Although no approach or technique will result in an accurate forecast, a combination of qualitative and quantitative approaches can be used to great effect by bringing together expert judgements and predictive models.

Qualitative methods

Imagine you were asked to forecast the outcome of a forthcoming football match. Simply looking at the teams' performance over the last few weeks and extrapolating it is unlikely to yield the right result. Like many business decisions the outcome will depend on many other factors. In this case the strength of the opposition, their recent form, injuries to players on both sides, the match location and even the weather will have an influence on the outcome. A qualitative approach involves collecting and appraising judgements, options, even best guesses, as well as past performance, from 'experts' to make a prediction. There are several ways this can be done: a panel approach, the Delphi method and scenario planning.

Panel approach

Just as panels of football pundits gather to speculate about likely outcomes, so too do politicians, business leaders, stock market analysts, banks and airlines. The panel acts like a focus group, allowing everyone to talk openly and freely. Although there is the great advantage of several brains being better than one, it can be difficult to reach a consensus, or sometimes the views of the loudest, or highest status, may emerge (the bandwagon effect). Although more reliable than one person's views, the panel approach still has the weakness that everybody, even the experts, can get it wrong.

Delphi method

Perhaps the best-known approach to generating forecasts using experts is the Delphi method.[1] This is a more formal method, which attempts to reduce the influences from the procedures of face-to-face meetings. It employs a questionnaire, emailed or posted to the experts. The replies are analysed and summarised and returned, anonymously, to all the experts. The experts are then asked to reconsider their original response in the light of the replies and arguments put forward by the other experts. This process is repeated several more times to conclude either with a consensus, or at least a narrower range of decisions. One refinement of this approach is to allocate weightings to the individuals and their suggestions based on, for example, their experience, their past success in forecasting, or other people's views of their abilities. The obvious problems associated with this method include constructing an appropriate questionnaire, selecting an appropriate panel of experts and trying to deal with their inherent biases.

Scenario planning

A method for dealing with situations of even greater uncertainty is scenario planning. This is usually applied to long-range forecasting, again using a panel. The panel members are usually asked to devise a range of future scenarios. Each scenario can then be discussed and the inherent risks considered. Unlike the Delphi method, scenario planning is not necessarily concerned with arriving at a consensus; it looks at the possible range of options, putting plans in place to try to avoid the ones that are least desired and taking action to follow the most desired.

Quantitative methods

There are two main approaches to quantitative forecasting – time series analysis and causal modelling techniques:

1. Time series analysis examines the pattern of past behaviour of a single phenomenon over time, taking into account reasons for variation in the trend in order to use the analysis to forecast the phenomenon's future behaviour.
2. Causal modelling is an approach that describes and evaluates the complex cause–effect relationships between the key variables (such as those in Figure 4.11).

Time series analysis

Simple time series plot a variable over time, then by removing underlying variations with assignable causes use extrapolation techniques to predict future behaviour. The key weakness with this approach is that it simply looks at past behaviour to predict the future, ignoring causal variables that are taken into account in other methods such as causal modelling or qualitative techniques. For example, suppose a company is attempting to predict the future sales of a product. The past three years' sales, quarter by quarter, are shown in Figure 4.13(a). This series of past sales may be analysed to indicate future sales. For instance, underlying the series might be a linear upward trend in sales. If this is taken out of the data, as in Figure 4.13(b), we are left with a cyclical seasonal variation. The mean deviation of each quarter from the trend line can now be taken out, to give the average seasonality deviation. What remains is the random variation about the trends and seasonality lines, Figure 4.13(c). Future sales may now be predicted as lying within a band about a projection of the trend, plus the seasonality. The width of the band will be a function of the degree of random variation.

Forecasting unassignable variations

The random variations that remain after taking out trend and seasonal effects are without any known or assignable cause. This does not mean that they do not have a cause, however, just that we do not know what it is. Nevertheless, some attempt can be made to forecast it, if only on the basis that future events will, in some way, be based on past events. We will examine

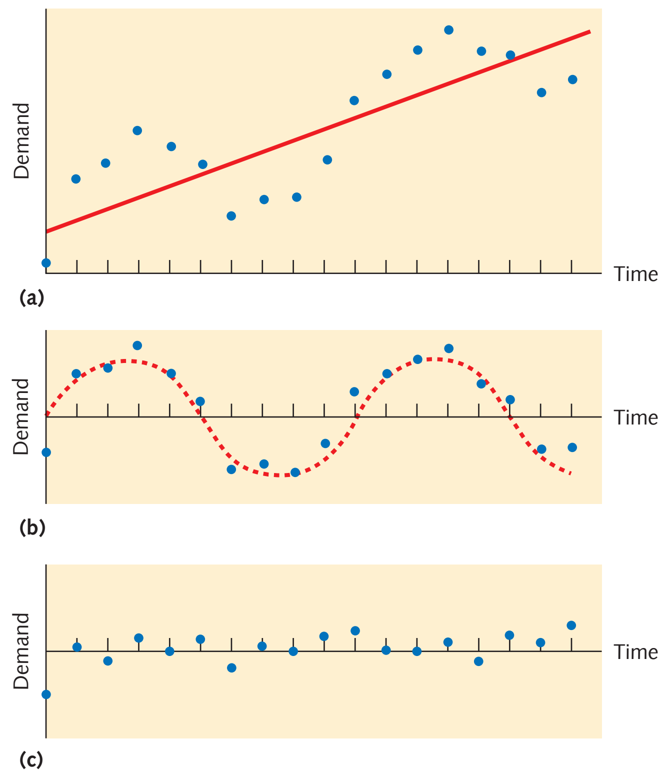


Figure 4.13 Time series analysis with (a) trend, (b) seasonality and (c) random variation

two of the more common approaches to forecasting, which are based on projecting forward from past behaviour. These are:

- moving-average forecasting;
- exponentially smoothed forecasting.

The moving-average approach to forecasting takes the previous *n* periods' actual demand figures, calculates the average demand over the *n* periods and uses this average as a forecast for the next period's demand. Any data older than the *n* periods play no part in the next period's forecast. The value of *n* can be set at any level, but is usually in the range 4 to 7.

Worked example

Eurospeed parcels

Table 4.4 shows the weekly demand for Eurospeed, a Europe-wide parcel delivery company. It measures demand, on a weekly basis, in terms of the number of parcels that it is given to deliver (irrespective of the size of each parcel). Each week, the next week's demand is forecast by taking the moving average of the previous four weeks' actual demand. Thus if the forecast demand for week t is F_t and the actual demand for week t is A_t, then:

$$F_t = \frac{1}{4}(A_{t-4} + A_t + A_{t-2} + A_{t-1})$$

For example, the forecast for week 35:

$$F_{35} = (72.5 + 66.7 + 68.3 + 67.0)/4$$
$$= 68.8$$

Table 4.4 Moving-average forecast calculated over a four-week period

Week	*Actual demand (thousands)*	*Forecast*
20	63.3	
21	62.5	
22	67.8	
23	66.0	
24	67.2	64.9
25	69.9	65.9
26	65.6	67.7
27	71.1	66.3
28	68.8	67.3
29	68.4	68.9
30	70.3	68.5
31	72.5	69.7
32	66.7	70.0
33	68.3	69.5
34	67.0	69.5
35		68.8

Exponential smoothing

There are two significant drawbacks to the moving-average approach to forecasting. First, in its basic form, it gives equal weight to all the previous *n* periods that are used in the calculations (although this can be overcome by assigning different weights to each of the *n* periods). Second, and more important, it does not use data from beyond the *n* periods over which the moving average is calculated. Both these problems are overcome by exponential smoothing, which is also somewhat easier to calculate. The exponential-smoothing approach forecasts demand in the next period by taking into account the actual demand in the current period and the forecast that was previously made for the current period. It does so according to the formula:

$$F_t = \alpha A_{t-1} + (1 - x)F_{t-1}$$

where α = the smoothing constant.

The smoothing constant α is, in effect, the weight that is given to the last (and therefore assumed to be most important) piece of information available to the forecaster. However, the other expression in the formula includes the forecast for the current period, which included the previous period's actual demand, and so on. In this way all previous data have a (diminishing) effect on the next forecast.

Table 4.5 shows the data for Eurospeed's parcels forecasts using this exponential-smoothing method, where $\alpha = 0.2$. For example, the forecast for week 35 is:

$$F_{35} = (0.2 \times 67.0) + (0.8 \times 68.3) = 68.04$$

The value of α governs the balance between the responsiveness of the forecasts to changes in demand, and the stability of the forecasts. The closer α is to 0 the more forecasts will be dampened by previous forecasts (not very sensitive but stable). Figure 4.14 shows the Eurospeed volume data plotted for a four-week moving average, exponential smoothing with $\alpha = 0.2$ and exponential smoothing with $\alpha = 0.3$.

Seasonality in forecasting

Most organisations experience seasonal patterns in their demand. Sometimes the causes of seasonality are climatic (holidays), sometimes festive (gift purchases), sometimes financial (tax processing), or social, or political. For most of us, we typically think of seasonality in annual terms. However, in forecasting the term is used to describe any regularly repeating changes in demand (quarterly, monthly, weekly, daily or hourly). For example, utility companies may experience large annual seasonality, but will also face seasonal patterns over the week and across each day. Similarly, repeating patterns of the daily and weekly demand patterns of a supermarket will fluctuate, with some degree of predictability. Demand might be low in the morning, higher in the afternoon, with peaks at lunchtime and after work in the evening. Demand might be low on Monday and Tuesday, build up during the latter part of the week and reach a peak on Friday and Saturday. A popular technique for incorporating seasonality in forecasting is the multiplicative seasonal model, where seasonal factors are multiplied by an estimate of average demand to generate a seasonal forecast. For simplicity,

Table 4.5 **Exponentially smoothed forecast calculated with smoothing constant $\alpha = 0.2$**

Week (t)	*Actual demand (thousands) (A)*	*Forecast (F_1 = can't do this)*
20	63.3	60.00
21	62.5	60.66
22	67.8	60.03
23	66.0	61.58
24	67.2	62.83
25	69.9	63.70
26	65.6	64.94
27	71.1	65.07
28	68.8	66.28
29	68.4	66.78
30	70.3	67.12
31	72.5	67.75
32	66.7	68.70
33	68.3	68.30
34	67.0	68.30
35		68.04

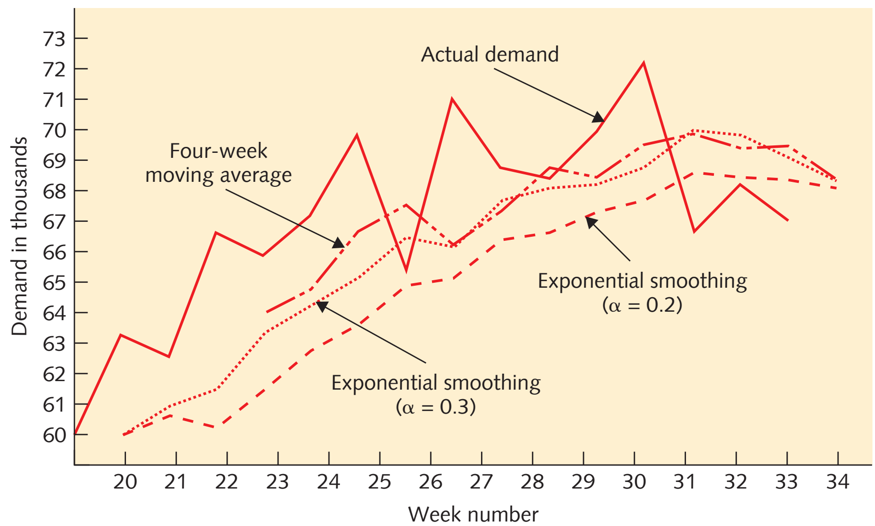


Figure 4.14 **A comparison of a moving-average forecast and exponential smoothing with the smoothing constant $\alpha = 0.2$ and 0.3**

here we assume that there is no other trend in the data, apart from seasonality. The seasonal model involves the following five steps:

1. Find the average demand for each 'season' by summing the demand for that season and dividing by the number of seasons available. For example, if in March we have had sales of 80, 75 and 100 over the last three years, average March demand equals $(80 + 75 + 100) / 3 = 85$.

2. Calculate average demand over all 'seasons' by dividing total average demand by the number of seasons. For example, if total average annual demand is 1320 and there are 12 seasons (months), average demand equals $1320/12 = 110$.

OPERATIONS PRINCIPLE
In demand forecasting, 'seasonality' refers to any repeating pattern of demand – annual, quarterly, monthly, weekly, daily or even hourly.

3. Compute seasonal index by dividing average season demand (step 1) over total average demand (step 2). For example, March seasonal index equals 85/110 = 0.77.
4. Estimate next time period's (in this case, annual) total demand using one or more of the qualitative or quantitative methods described in this section.
5. Divide this estimate by the number of seasons (in this case, 12) and multiply by the seasonal index to provide a seasonal forecast.

Worked example

Phoenix Consulting expected to have an annual demand for 7,500 hours of supply chain strategy consulting in 2018. Using the multiplicative seasonal model, we can forecast demand for June, July and August of that year (see Table 4.6).

Causal models

Causal models often employ complex techniques to understand the strength of relationships between the network of variables and the impact they have on each other. Simple regression models try to determine the 'best fit' expression between two variables. For example, suppose an ice-cream company is trying to forecast its future sales. After examining previous demand, it figures that the main influence on demand at the factory is the average temperature of the previous week. To understand this relationship, the company plots demand against the previous week's temperatures. This is shown in Figure 4.15. Using this graph, the company can make a reasonable prediction of demand, once the average temperature is known, provided that the other conditions prevailing in the market are reasonably stable. If they are not, then these other factors that have an influence on demand will need to be included in the regression model, which becomes increasingly complex.

These more complex networks comprise many variables and relationships, each with their own set of assumptions and limitations. Many techniques are available to help managers undertake this more complex modelling. They can be used to assess the importance of each of

Table 4.6 Hours of consulting sold

Month	*2015*	*2016*	*2017*	*Avg. of 3-yr demand*	*Avg. monthly demand*	*Seasonal index*
Jan	450	475	475	466.67	570.14	0.82
Feb	500	500	550	516.67	570.14	0.91
Mar	625	600	575	600.00	570.14	1.05
Apr	600	600	650	616.67	570.14	1.08
May	550	600	600	583.33	570.14	1.02
Jun	600	625	650	625.00	570.14	1.10
Jul	700	750	800	750.00	570.14	1.32
Aug	450	400	500	450.00	570.14	0.79
Sep	500	450	450	466.67	570.14	0.82
Oct	550	500	525	525.00	570.14	0.92
Nov	650	600	650	633.33	570.14	1.11
Dec	600	600	625	608.33	570.14	1.07
Total average annual demand				6841.67		

June 2018 forecast = (7,500/12) × 1.10 = 687.50
July 2018 forecast = (7,500/12) × 1.32 = 825.00
August 2018 forecast = (7,500/12) × 0.79 = 493.75

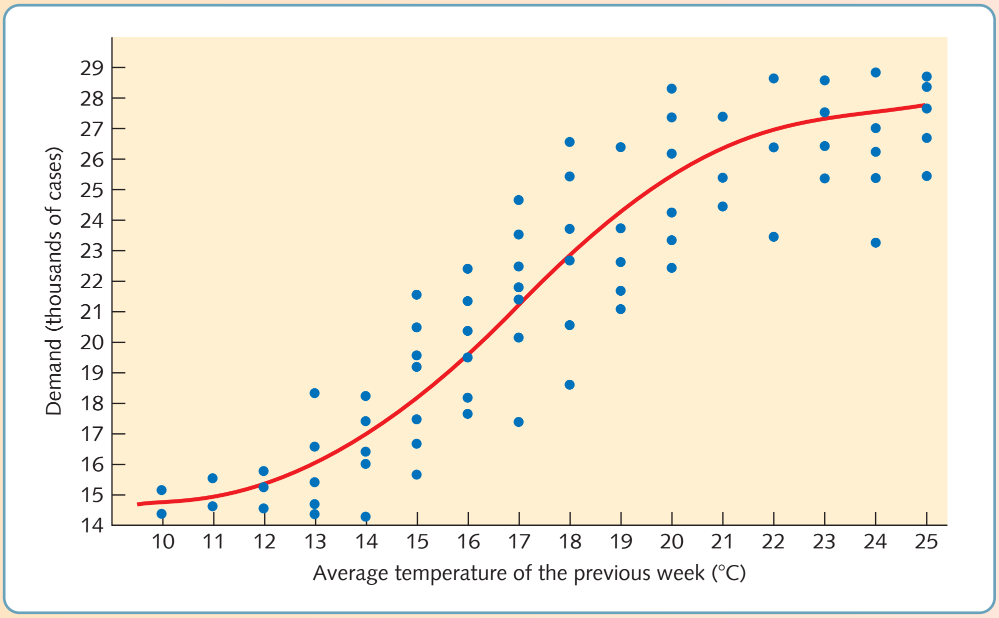


Figure 4.15 Regression line showing the relationship between the previous week's average temperature and demand

the factors, understand the network of interrelationships between factors and introduce feedback data into the model. However, such sophisticated models are beyond the scope of this text.

The performance of forecasting models

Forecasting models are widely used in management decision-making, and indeed most decisions require a forecast of some kind, yet the performance of this type of model is far from impressive. Hogarth and Makridakis,[2] in a comprehensive review of the applied management and finance literature, show that the record of forecasters using both judgement and sophisticated mathematical methods is not good. What they do suggest, however, is that certain forecasting techniques perform better under certain circumstances. In short-term forecasting there is considerable inertia in most economic and natural phenomena. Thus, the present states of any variables are predictive of the short-term future (i.e. 3 months or less). So, simple mechanistic methods, such as those used in time series forecasts, can often make accurate short-term forecasts and even out-perform more theoretically elegant and elaborate approaches used in econometric forecasting.[3]

However, long-term forecasting methods, although difficult to judge because of the time lapse between the forecast and the event, do seem to be more amenable to an objective causal approach. In a comparative study of long-term market forecasting methods, Armstrong and Grohman[4] conclude that econometric methods offer more accurate long-range forecasts than do expert opinion or time series analysis, and that the superiority of objective causal methods improves as the time horizon increases.

Notes on chapter supplement

1 Linstone, H.A. and Turoof, M. (eds) (1975) *The Delphi Method: Techniques and Applications*, Addison-Wesley.

2 Hogarth, R.M. and Makridakis, S. (1981) 'Forecasting and planning: An evaluation', *Management Science*, (27) 2, pp. 115–138.

3 Hogarth, R.M. and Makridakis, S. (1981) *op. cit*.

4 Armstrong, J.S. and Grohman, M.C. (1972) 'A comparative study of methods for long-range market forecasting', *Management Science,* (19) 2, pp. 211–221.

Taking it further

Hendry, D., Castle, J. and Clements, M *(2019) Forecasting: An essential introduction, Yale University Press.* *As it says, an introduction*.

Hyndman, R.J. and Athanasopoulos, G. *(2018) Forecasting: Principles and practice, 2nd edition, OTexts.* *A textbook that provides a comprehensive introduction to forecasting methods*.

Morlidge, S. and Player, S. *(2010) Future Ready: How to master business forecasting, John Wiley & Sons.* *A good, not too technical, treatment*.

Silver, N. *(2013) The Signal and the Noise: The art and science of prediction, Penguin.* *Nate Silver is probably the best-known forecaster in the world, and for good reason (he accurately predicted the results of every single state in the 2012 US election). A very readable book.*

Tetlock, P. and Gardner, D. *(2016) Superforecasting: The art and science of prediction, Random House Books.* *Readable and sensible*.

5 Process design 1 – Positioning

Introduction

In Chapter 1 we described a 'process' as an arrangement of resources and activities that transform inputs into outputs that satisfy (internal or external) customer needs. Processes are everywhere. They are the building blocks of all operations, and their design will affect the performance of the whole operation and, eventually, the contribution it makes to its supply network. No one, in any function or part of the business, can fully contribute to its competitiveness if the processes in which they work are poorly designed and ineffective. It is not surprising then that process design has become such a popular topic in the management press and among consultants. This chapter is the first of two that examine the design of processes. To understand the difference between this chapter (positioning) and the following one (analysis), go back to our definition of a process: 'an arrangement of resources and activities'. This chapter is primarily concerned with the resources in processes and, more specifically, how process resources normally reflect the volume and variety requirements placed on them. The next chapter examines the activities within processes, and how they can be analysed in order to understand better how they will operate and, therefore, how their performance could be improved. Figure 5.1 shows the position of the ideas described in this chapter within the general model of operations management.

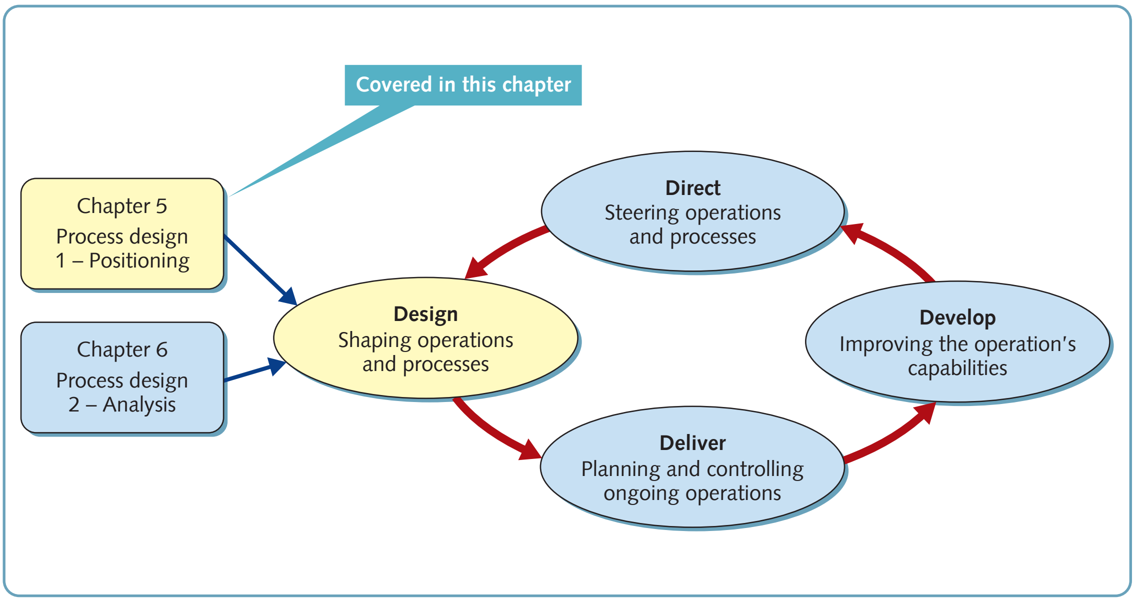
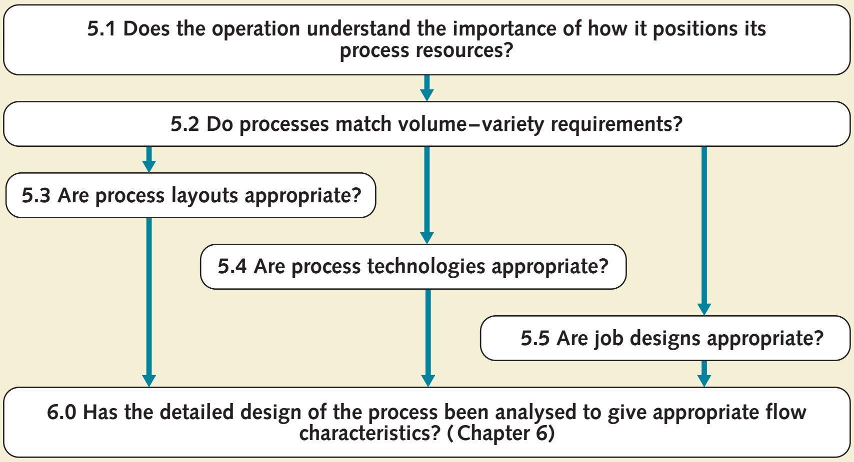


Figure 5.1 Process design – positioning is concerned with conceiving the nature of the resources that make up the process so that they are appropriate for their volume–variety position

EXECUTIVE SUMMARY



5.1 Does the operation understand the importance of how it positions its process resources?

Process design is concerned with conceiving the nature of the resources that make up the process and with their detailed workings. The first task of process design is to conceive the nature of the resources that make up the process and how they are arranged. Without appropriate resources it is difficult (maybe impossible) for the process ever to operate as effectively as it could do. A useful way to understand this is to position the process in terms of its volume and variety characteristics. Only after this stage should the second task (of conceiving the detailed workings of the process) be attempted.

5.2 Do processes match volume–variety requirements?

Volume and variety are particularly influential in the design of processes. They also tend to go together in an inverse relationship. High-variety processes are normally low in volume and vice versa. So processes can be positioned on the spectrum between low volume and high variety and high volume and low variety. At different points on this spectrum, processes can be described as distinct process 'types'. Different terms are used in manufacturing and service to identify these types. Working from low volume and high variety towards high volume and low variety, the process types are: project processes; jobbing processes; batch processes; mass processes; and continuous processes. Working in the same sequence the service types are known as: professional services; service shops; and mass services. Whatever terminology is used, the overall design of the process must fit its volume–variety position. This is usually summarised in the form of the 'product–process' matrix.

5.3 Are process layouts appropriate?

There are different ways in which the different resources within a process (people and technology) can be arranged relative to each other. But however this is done it should reflect the process's volume–variety position. Again, there are pure 'types' of layout that correspond with the different volume–variety positions. These are: fixed-position layout; functional layout; cell layout; and product layout. Many layouts are hybrids of these pure types, but the type chosen is influenced by the volume and variety characteristics of the process.

5.4 Are process technologies appropriate?

Process technologies are the machine's equipment and devices that help processes to transform materials, information and customers. They are different to the product technology that is embedded within the product or service itself. Again, process technology should reflect volume and variety. In particular, the degree of automation in the technology, the scale and/or scalability of the technology and the coupling and/or connectivity of the technology should be appropriate to volume and variety. Generally, low volume and high variety requires relatively unautomated, general-purpose, small-scale and flexible technologies. By contrast, high-volume and low-variety processes require automated, dedicated and large-scale technologies that are sometimes relatively inflexible.

5.5 Are job designs appropriate?

Job design is about how people carry out their tasks within a process. It is particularly important because it governs people's expectations and perceptions of their contribution to the organisation, as well as being a major factor in shaping the culture of the organisation. Some aspects of job design are common to all processes, irrespective of their volume and variety position. These are such things as ensuring the safety of everyone affected by the process, ensuring a firm ethical stance and upholding an appropriate work–life balance. However, other aspects of job design are influenced by volume and variety – in particular, the extent of division of labour, the degree to which jobs are defined and the way in which job commitment is encouraged. Broadly, high-variety and low-volume processes require broad, relatively undefined jobs with decision-making discretion. Such jobs tend to have intrinsic job commitment. By contrast, high-volume and low-variety processes tend to require jobs that are relatively narrow in scope and closely defined with relatively little decision-making discretion. This usually means some deliberative action is needed in the design of the job (such as job enrichment) in order to help maintain commitment.

5.1 Diagnostic question: *Does the operation understand the importance of how it positions its process resources?*

'Positioning' the resources in processes is the first step in making sure that they are well designed. To 'design' is to conceive the looks, arrangement and workings of something before it is constructed. In that sense, design is a conceptual exercise. Yet it is also one that must deliver a solution that will work in practice. Design is also an activity that can be approached at different levels of detail. One may envisage the general shape and intention of something before getting down to defining its details. However, it is often only through getting to grips with the detail of a design that the feasibility of its overall shape can be assessed. So it is with designing processes. First, one must consider the overall shape and nature of the resources that make up the process. Second, one must analyse the details of the activities within the process in order to ensure that it fulfils its objectives effectively. But don't think of this as a simple sequential process. There may be aspects concerned with the broad positioning of the process that will need to be modified following its more detailed analysis.

It is worth noting that an operation's resources are more than what is shown directly on its balance sheet. They include not only a business's technology, equipment and facilities that form its physical fabric, but also its people (including their skills, enthusiasm and creativity), and its intangible elements (such as knowledge, supply and customer relationships, culture, etc.) together with their intrinsic capabilities. Note that an operation's assets also include the resources that it can conveniently 'get access to' (usually through its supply network). This means that an operation need not necessarily own the resources that it uses. It could have supply agreements that allow it to access resources as and when it has the need. In effect, an operation's ability to access resources is a resource in itself.

Process positioning and process analysis are related

Although we have separated process design into two chapters, the positioning of resources and the detailed analysis of processes are not independent activities; they are simply a way of thinking about the process design activity. Moreover, process positioning and analysis are closely related and impact on each other. They can be both mutually supportive and mutually limiting. The choice and positioning of resources will constrain what an operation's processes are capable of doing. Technology, people or systems with inappropriate abilities or skills will restrict a process's abilities. Just as important, but sometimes underappreciated, is how the way one manages a process's detailed design can enhance its resources. For example, a highly divided, repetitive process design could limit the accumulation of learning and hence operations capabilities. Ideally, processes should not simply 'do what they do' without learning from what they do. Processes are really the best source of learning that operations managers will ever have. In any medium to large business, thousands of individual transactions take place every minute. Every transaction is an opportunity to learn. This idea, and the purpose of this and the next chapter, is illustrated in Figure 5.2.

> **OPERATIONS PRINCIPLE**
> ***Defining process resources (process positioning) and defining process activities (process analysis) are strongly related and ideally should be considered alongside one another.***

5.2 Diagnostic question: *Do processes match volume–variety requirements?*

Two factors are particularly important in process design: these are the volume and variety of the products and services that it processes. Moreover, volume and variety are related, in so much as low-volume operations processes often have a high variety of products and services, and

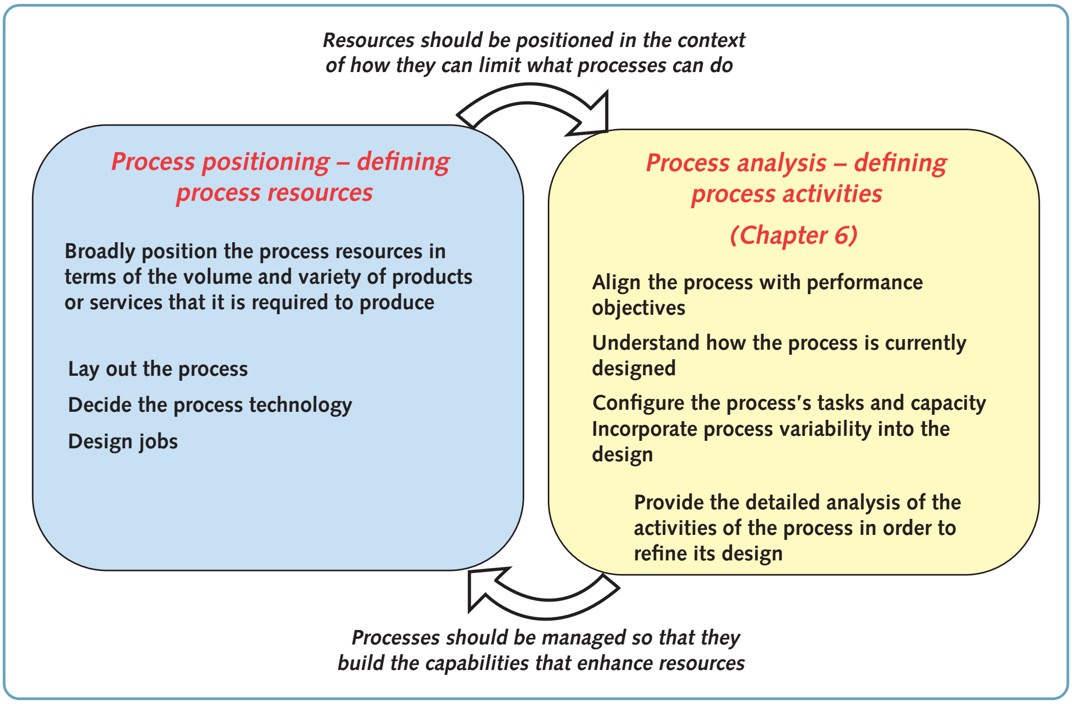


Figure 5.2 **Process design is treated in two parts: positioning, which sets the broad characteristics of the resources in the process; and analysis, which refines the details of the design by examining the activities within the process**

high-volume operations processes often have a narrow variety of products and services. So, we can position processes on a continuum from those that operate under conditions of low volume and high variety, through to those that operate under conditions of high volume and low variety. The volume–variety position of a process influences almost every aspect of its design. Processes with different volume–variety positions will be arranged in different ways, have different flow characteristics and have different technology and jobs. So, a first step in process design is to understand how volume and variety shape process characteristics, and to check whether processes have been configured in a manner that is appropriate for their volume–variety position.

> OPERATIONS PRINCIPLE
> *The design of any process should be governed by the volume and variety it is required to produce.*

The 'product–process' matrix

The most common method of illustrating the relationship between a process's volume–variety position and its design characteristics is shown in Figure 5.3. Often called the 'product–process' matrix, it can in fact be used for any type of process, whether producing products or services. The underlying idea of the product–process matrix is that many of the more important elements of process design are strongly related to the volume–variety position of the process. So, for any process, the tasks that it undertakes, the flow of items through the process, the layout of its resources, the technology it uses and the design of jobs, are all strongly influenced by its volume–variety position. This means that most processes should lie close to the diagonal of the matrix that represents the 'fit' between the process and its volume–variety position. This is called the 'natural' diagonal.

> OPERATIONS PRINCIPLE
> *Process types indicate the position of processes on the volume–variety spectrum.*

Process types

Processes that inhabit different points on the diagonal of the product–process matrix are sometimes referred to as 'process types'. Each process type implies differences in the set of tasks

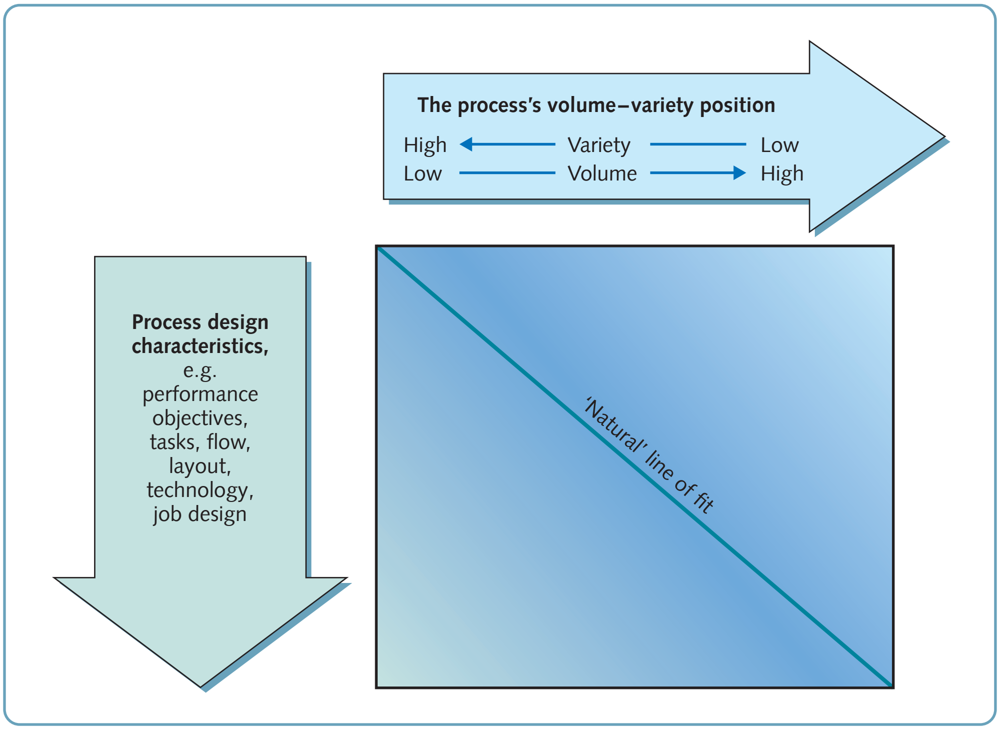


Figure 5.3 **The elements of process design are strongly influenced by the volume–variety requirements placed on the process**

Case example

Dishang and Sands Films – at opposite ends of the volume–variety spectrum[1]

Making clothes must have been one of the very first 'production' tasks carried out by early humans, and it is still an important industrial sector. With the world garment market worth around €1.3 trillion, it is estimated as employing up to 75 million people.[2] Of all garment-producing countries, China has long been seen as the master of clothing manufacturing, and its largest producer and exporter of apparel and textiles is the Dishang Group. Founded in 1993, Dishang has an annual turnover of $1.5bn, producing garments for such well-known brands as Zara, Matalan and Adidas. But, although China is still a leader in terms of technical expertise and production efficiency, when its labour costs increased, Dishang, like other producers, expanded to set up operations in Cambodia, Myanmar and Bangladesh. They are a huge and sophisticated enterprise, manufacturing over 73 million garments per year from 80 wholly-owned factories across 12 global locations. Their modern operations include automated technology, full on-line and end-of-line quality control systems and the use of acceptable quality limit (AQL) inspection levels according to the requirements of each customer. At Dishang's headquarters, customers can peruse 50,000 fabrics in the company's large digital library. Once customers have chosen a pattern they like, they can upload a picture into an internal system and be shown similar styles. Dishang's Chairman, Lihua Zhu, believes that the Group's success is down to three factors: '*Firstly our volume strength is very important. Due to the size of the group we can secure the more competitive prices for our customers (when purchasing raw materials and components). Secondly we offer in-house design expertise and have factories and internal teams that specialise in different products, meaning brands and retailers can come to us for everything. And thirdly, we approach international markets differently with our own offices, which saves costs by cutting out the middle man.*'[3]

One certainly would not see such high volume at Sands Films' costume-making workshop. Every film or television programme that is set in any period, other than the present day, needs costumes for its actors. And most films have a lot of characters, so that means a lot of costumes. Sands Films in London has a well-established and permanent garment workshop. It is what we will describe later in the chapter as a typical 'jobbing' process. Sands Films provides a wide range of wardrobe and costume services. Its customers are the film, stage and TV production companies, each of which have different requirements and time constraints. And because each project is different and has different requirements, the workshop's jobs go from making a single simple outfit to providing a wide variety of specially designed costumes and facilities over an extended production period. The facilities include most normal tailoring processes such as cutting, dyeing and printing, and varied specialist services such as corset- and crinoline-making, as well as millinery (hats). During the design and making process, actors often visit the workshop, which has been called an 'Aladdin's cave' of theatrical costumes. The workshop is where actors come face to face with their character for the first time. Making a costume can only start once a project has been approved and a costume designer appointed, although discussions with the workshop may have started prior to that. When the budget and the timing have been agreed, the designer can start to present ideas and finished designs to the workshop. And although the processes in the workshop are well established, each costume requires different skills and so has different routes through the stages.

BBC Films/Focus Features/Kobal/Shutterstock

performed by the process and in the way materials information, or customers, flow through the process. Different terms are sometimes used to identify process types depending on whether they are predominantly manufacturing or service processes and there is some variation in how the names are used. This is especially so in service process types. It is not uncommon to find manufacturing terms used also to describe service processes. Perhaps most importantly, there is some degree of overlap between process types. The different process types are shown in Figure 5.4.

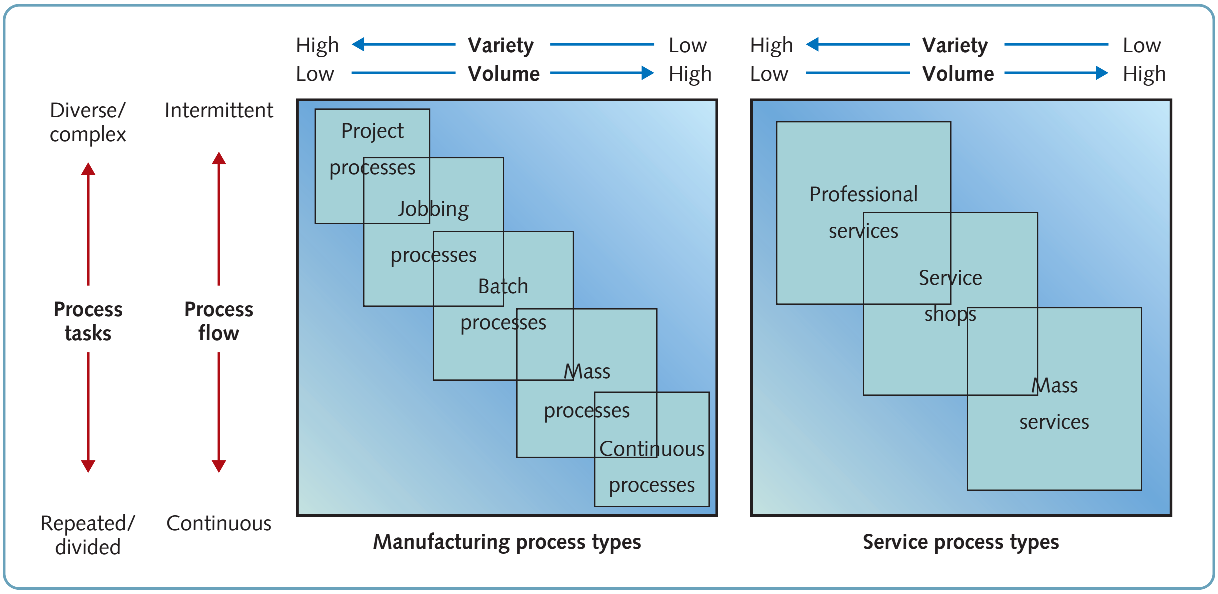


Figure 5.4 **Different process types imply different volume–variety characteristics for the process**

Project processes

Project processes are those that deal with discrete, usually highly customised products. Often the timescale of making the product is relatively long, as are the intervals between the completion of each product. The activities involved in the process may be ill-defined and uncertain, sometimes changing during the process itself. Examples include advertising agencies, shipbuilding, most construction companies, movie production companies, drilling oil wells and installing computer systems. Any process map for project processes will almost certainly be complex, partly because each unit of output is usually large with many activities occurring at the same time, and partly because the activities often involve significant discretion to act according to professional judgement. In fact, a process map for a whole project would be extremely complex, so rarely would a whole project be mapped, but small parts of it may be.

Jobbing processes

Jobbing processes also deal with very high variety and low volumes, but whereas in project processes each project has resources devoted more or less exclusively to it, in jobbing processes each 'product' has to share the operation's resources with many others. The process will work on a series of products but, although all the products will require the same kind of attention, each will differ in its exact needs. Examples of jobbing processes include many precision engineers such as specialist toolmakers, furniture restorers, 'made-to-measure' tailors and the printer who produces tickets for the local social event. Jobbing processes produce more and usually smaller items than project processes but, like project processes, the degree of repetition is low. Many jobs could be 'one-offs'. Again, any process map for a jobbing process could be relatively complex for similar reasons to project processes. Although jobbing processes sometimes involve considerable skill, they are usually more unpredictable than project processes.

Batch processes

Batch processes can look like jobbing processes, but without the degree of variety normally associated with jobbing. As the name implies, batch processes usually produce more than one 'product' at a time. So each part of the operation has periods when it is repeating itself, at least while the 'batch' is being processed. The size of the batch could be just two or three, in which case the batch process would differ little from jobbing, especially if each batch is a totally novel product. Conversely, if the batches are large, and especially if the products are familiar to the operation, batch processes can be fairly repetitive. Because of this, the batch type of process can be found over a wider range of volume and variety levels than other process types. Examples of batch processes include machine tool manufacturing, the production of some special gourmet frozen foods, the manufacture of most of the component parts that go into mass-produced assemblies such as automobiles, and the production of most clothing. Batch process maps may look straightforward, especially if different products take similar routes through the process with relatively standard activities being performed at each stage.

Mass processes

Mass processes produce in high volume, usually with narrow effective variety. A car manufacturing plant, for example, might produce several thousand variants of cars if every option of engine size, colour and equipment is taken into account. Yet its effective variety is low because the different variants do not affect the basic process of production. The activities in the car manufacturing plant, like all mass processes, are essentially repetitive and largely predictable. In addition to the car plant, examples of mass processes include consumer durables manufacturers, most food processes such as a frozen pizza manufacturer, beer bottling plants and CD production. Process maps for this type of process will be straightforward sequences of activities.

Continuous processes

Continuous processes are one step beyond mass processes in so much as they operate at even higher volume and often have even lower variety. Sometimes they are literally continuous in that their products are inseparable, being produced in an endless flow. Continuous processes are often associated with relatively inflexible, capital-intensive technologies with highly predictable flow. Examples of continuous processes include petrochemical refineries, electricity utilities, steel making and internet server farms. Like mass processes, process maps will show few elements of discretion, and although products may be stored during the process, the predominant characteristic of most continuous processes is a smooth flow from one part of the process to another.

Professional services

Professional services are high-variety, low-volume processes, where customers may spend a considerable time in the service process. Such services usually provide high levels of customisation, so contact staff are given considerable discretion. They tend to be people-based rather

Case example

Legal and General's modular housing process[4]

Legal & General (L&G) is not the type of company one would expect to be building houses. One of the UK's leading financial services groups, who have invested over £19 billion in projects including homebuilding, urban regeneration and clean energy, they became involved in building modular homes. Modular construction of housing is more like the way you would expect an automobile to be made. Modules are made 'off-site' in a factory then transported to the building site. As some modular construction proponents pointed out, once all cars were hand-built, but now cars are assembled in a factory. Rosie Toogood, the Chief Executive, and champion of the modern manufacturing and construction approach, came from aerospace company Rolls-Royce, and Stuart Lord, the Operations Director, spent his prior career in the automotive industry.

Starting a modular homes business from scratch required the considerable investment that a major financial group like L&G could provide. By retaining quality construction but using modern processes L&G called their approach 'everything new, but nothing new'. What this meant was an assembly line that had four giant computer-operated cutting and milling machines, and four smaller ones, all of which were capable of cutting timber panels to more precise levels of precision than could normally be achieved on a conventional building site. The finished modules, fitted with wiring and plumbing, decorated, carpeted and fitted out with kitchens and bathrooms, could then be loaded onto a lorry and delivered to sites. A home could consist of a single module or several combined. This approach meant digitally modelling every millimetre of every home before production started, standardising and simplifying processes for efficient, high-quality production. Just as important, it involved embracing continuous improvement – the approach that almost halved the time to deliver a completed house. By standardising and simplifying processes, they could drive up quality and productivity, thus driving down costs. Also, the adoption of mass processing was influenced by increasingly tough energy targets. The process used less water than conventional building methods and reduced site-generated waste. This, in turn, required fewer skips and provided a tidier, safer site. A report by the Ellen MacArthur Foundation highlighted several potential environmental benefits from off-site construction, including more energy-efficient homes.[5]

Shutterstock

than equipment-based, with emphasis placed on the process (how the service is delivered) as much as the 'product' (what is delivered). Examples include management consultants, lawyers' practices, architects, doctors' surgeries, auditors, health and safety inspectors and some computer field-service operations. Where process maps are used, they are likely to be drawn predominantly at a high level. Consultants, for example, frequently use a predetermined set of broad stages, starting with understanding the real nature of the problem through to the implementation of their recommended solutions. This high-level process map guides the nature and sequence of the consultants' activities.

Service shops

Service shops are characterised by levels of customer contact, customisation, volume of customers and staff discretion that position them between the extremes of professional and mass services (see below). Service is provided via a mix of front- and back-office activities. Service shops include banks, high-street shops, holiday tour operators, car rental companies, schools, most restaurants, hotels and travel agents. For example, an equipment hire and sales organisation may have a range of equipment displayed in front-office outlets, while back-office operations look after purchasing and administration. The front-office staff have some technical training and can advise customers during the process of selling the product. Essentially, the customer is buying a fairly standardised 'product' but will be influenced by the process of the sale, which is customised to the individual customer's needs.

Mass services

Mass services have many customer transactions and little customisation. Such services are often predominantly equipment based and 'product' oriented, with most value added in the back office, sometimes with comparatively little judgement needed by front-office staff, who may have a closely defined job and follow set procedures. Mass services include supermarkets, a national rail network, an airport and many call centres. For example, airlines move a large number of passengers on their networks. Passengers pick a journey from the range offered. The airline can advise passengers on the quickest or cheapest way to get from A to B, but they cannot 'customise' the service by putting on special flights just for them.

Moving off the natural diagonal

A process lying on the natural diagonal of the matrix shown in Figure 5.3 will normally have lower operating costs than one with the same volume–variety position that lies off the diagonal. This is because the diagonal represents the most appropriate process design for any volume–variety position. Processes that are on the right of the 'natural' diagonal would normally be associated with lower volumes and higher variety. This means that they are likely to be more flexible than seems to be warranted by their actual volume–variety position. That is, they are not taking advantage of their ability to standardise their activities. Because of this, their costs are likely to be higher than they would be with a process that was closer to the diagonal. Conversely, processes that are on the left of the diagonal have adopted a position that would normally be used for higher volume and lower variety processes. Processes will therefore be 'over-standardised' and probably too inflexible for their volume–variety position. This lack of flexibility can also lead to high costs because the process will not be able to change from one activity to another as readily as a more flexible process. One note of caution regarding this idea: although logically coherent, it is a conceptual model rather than something that can be 'scaled'. Although it is intuitively obvious that deviating from the diagonal increases costs, the precise amount by which costs will increase is very difficult to determine. Nevertheless, a first step in examining the design of an existing process is to check if it is on the natural diagonal of the

OPERATIONS PRINCIPLE

Moving off the 'natural diagonal' of the product–process matrix will incur excess costs.

product–process matrix. The volume–variety position of the process may have changed without any corresponding change in its design. Alternatively, design changes may have been introduced without considering their suitability for the process's volume–variety position.

Layout, technology and design

If movement down the natural diagonal of the product–process matrix changes the nature of a process, then the key elements of its design will also change. At this broad level, these 'key elements' are the two 'ingredients' that make up process resources (technology and people) and the way in which these ingredients are arranged within the process relative to each other. This latter aspect is usually called 'layout'. In the remainder of the chapter, we start by discussing layout and then the design decisions that relate to process technology and the jobs that the people within the process undertake.

Worked example

Water meters

The 'meter installation' unit of a water utility company installed and repaired water meters. The nature of each installation job could vary significantly because the metering requirements of each customer varied and because meters had to be fitted into very different water pipe systems. When a customer requested an installation or repair, a supervisor would survey the customer's water system and transfer the results of the survey to the installation team of skilled plumbers. An appointment would then be made for a plumber to visit the customer's location and install or repair the meter on the agreed appointment date. The company decided to install for free a new 'standard' remote-reading meter that would replace the wide range of existing meters and could be read automatically using the customer's telephone line. This would save meter-reading costs. It also meant a significant increase in work for the unit and more skilled plumbing staff were recruited. The new meter was designed to make installation easier by including universal quick-fit joints that reduced pipe cutting and jointing during installation. As a pilot, it was also decided to prioritise those customers with the oldest meters and conduct trials of how the new meter worked in practice. All other aspects of the installation process were left as they were.

The pilot was not a success. Customers with older meters were distributed throughout the company's area, so staff could not service several customers in one area and had to travel relatively long distances between customers. Also, because customers had not initiated the visit themselves, they were more likely to have forgotten the appointment, in which case plumbers had to return to their base and try to find other work to do. The costs of installation were proving to be far higher than forecast and the plumbers were frustrated at the waste of their time and the now relatively standardised installation job. The company decided to change its process. Rather than replace the oldest meters that were spread around its region, it targeted smaller geographic areas to limit travelling time. It also cut out the survey stage of the process because, using the new meter, 98 per cent of installations could be fitted in one visit, minimising disruption to the customer and the number of missed appointments. Just as significantly, fully qualified plumbers were often not needed, so installation could be performed by less-expensive labour.

This example is illustrated in Figure 5.5. The initial position of the installation process is at point A. The installation units were required to repair and install a wide variety of meters into a very wide variety of water systems. This needed a survey stage to assess the nature of the job and the use of skilled labour to cope with the complex tasks. The installation of the new type of meter changed the volume–variety position for the process by reducing the variety of the jobs tackled by the process and increasing the volume it had to cope with. However, the process was not changed. By choosing a wide geographic area to service, retaining the unnecessary survey stage and hiring over-skilled staff, the company was still defining itself as a high-variety, low-volume 'jobbing' process. The design of the process was appropriate for its old volume–variety position, but not the new one. In effect it had moved to point B in Figure 5.5. It was off the diagonal, with unnecessary flexibility and high operating costs. Redesigning the process to take advantage of the reduced variety and complexity of the job (position C in Figure 5.5) allowed installation to be performed far more efficiently.

▶

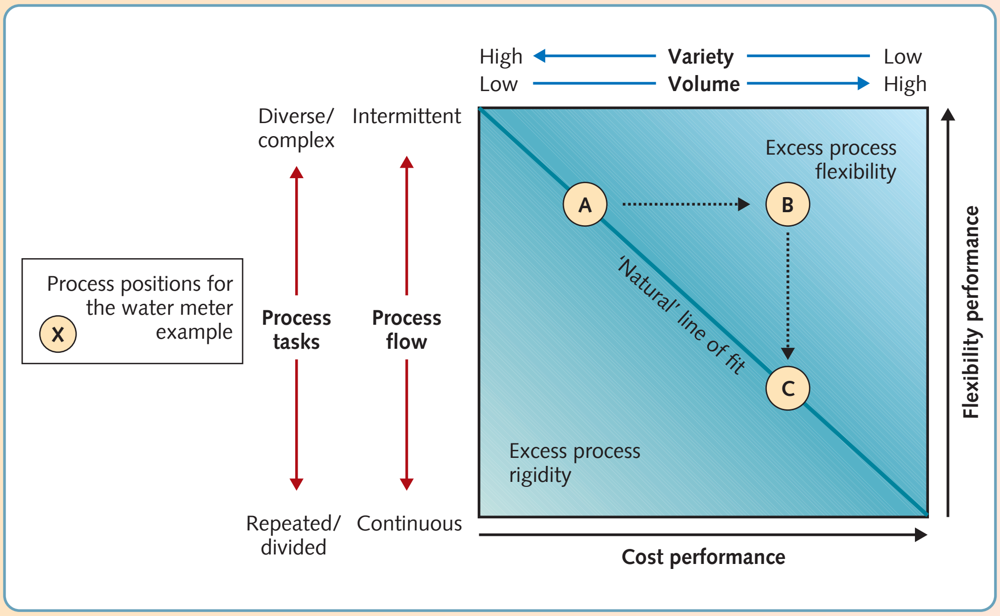


Figure 5.5 **A product–process matrix of the meter installation process showing different positions**

5.3 Diagnostic question: *Are process layouts appropriate?*

There is little point in having a well-sequenced process if in reality its activities are physically located in a way that involves excessive movement of materials, information or customers. Usually, the objective of the layout decision is to minimise movement, but, especially in an information-transforming process where distance is largely irrelevant, other criteria may dominate. For example, it may be more important to lay out processes such that similar activities or resources are grouped together. So, an international bank may group its foreign exchange dealers together to encourage communication and discussion between them, even though the 'trades' they make are processed in an entirely different location. Some high-visibility processes may fix their layout to emphasise the behaviour of the customers who are being processed.

Layout should reflect volume and variety

Again, the layout of a process is determined partly by its volume and variety characteristics. When volume is very low and variety is relatively high, 'flow' may not be a major issue. For example, in telecommunications satellite manufacture each product is different, and because products 'flow' through the operation very infrequently, it is not worth arranging facilities to minimise the flow of parts through the operation. With higher volume and lower variety, flow becomes a far more important issue. If variety is still high, however, an entirely flow-dominated arrangement is difficult because there will be different flow patterns. For example, a library will arrange its different categories of books and its other services partly to minimise the average distance its customers have to 'flow' through the operation. But, because its

> **OPERATIONS PRINCIPLE**
> ***Resources in low-volume, high-variety processes should be arranged to cope with irregular flow.***

> **OPERATIONS PRINCIPLE**
> ***Resources in high-volume, low-variety processes should be arranged to cope with smooth, regular flow.***

customers' needs vary, it will arrange its layout to satisfy the majority of its customers (but perhaps inconvenience a minority). When the variety of products or services reduces to the point where a distinct 'category' with similar requirements becomes evident but variety is still not small, appropriate resources could be grouped into a separate cell. When variety is relatively small and volume is high, flow can become regularised and resources can be positioned to address the (similar) needs of the products or services, as in a classic flow line.

Most practical layouts are derived from only four basic layout types that correspond to different positions on the volume–variety spectrum. These are illustrated diagrammatically in Figure 5.6 and are described here.

Fixed-position layout

Fixed-position layout is in some ways a contradiction in terms, since the transformed resources do not move between the transforming resources. Instead of materials, information or customers flowing through an operation, the recipient of the processing is stationary and the equipment, machinery, plant and people who do the processing move as necessary. This could be because the product or the recipient of the service is too large to be moved conveniently, or it might be too delicate to move, or perhaps it could object to being moved; for example:

- power generator construction – the product is too large to move;
- open-heart surgery – patients are too delicate to move;
- high-class restaurant – customers would object to being moved to where food is prepared.

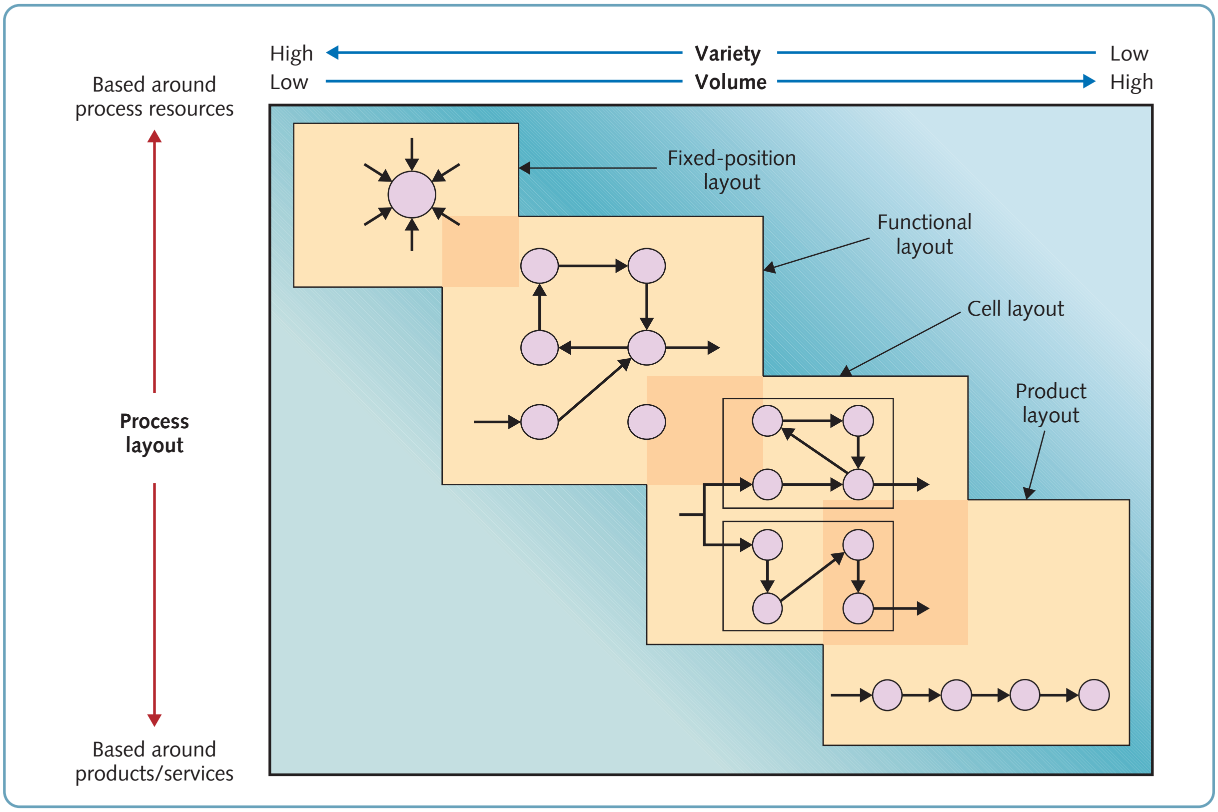


Figure 5.6 **Different process layouts are appropriate for different volume–variety combinations**

Case example

Reconciling quiet and interaction in laboratory layout[6]

The layout of scientific laboratories is rarely straightforward. Not only can different areas of a laboratory require very different service needs (temperature, extraction, lack of vibration, etc.) but also the types of work that all scientists engage in can have different needs. On one hand, there is active experimentation; on the other hand, there are times when quiet reflection is vital. Moreover, different individuals have different preferred working patterns. The conversations, discussions and, sometimes noisy, debate between some researchers can both irritate and distract other staff who prefer somewhere quiet to think and write up their work. Even in prestigious and high-profile research operations, this conflict can be difficult to reconcile. For example, some of the researchers working at the Francis Crick laboratory in central London complain that its open-plan layout, designed to encourage collaboration, makes it difficult to concentrate on their work. Some people like the background noise, while others prefer total silence, although many agree that the layout has been extremely successful in terms of promoting ad hoc meetings and has created new collaborations. Professor Alan Penn, who has been investigating how open-plan layouts

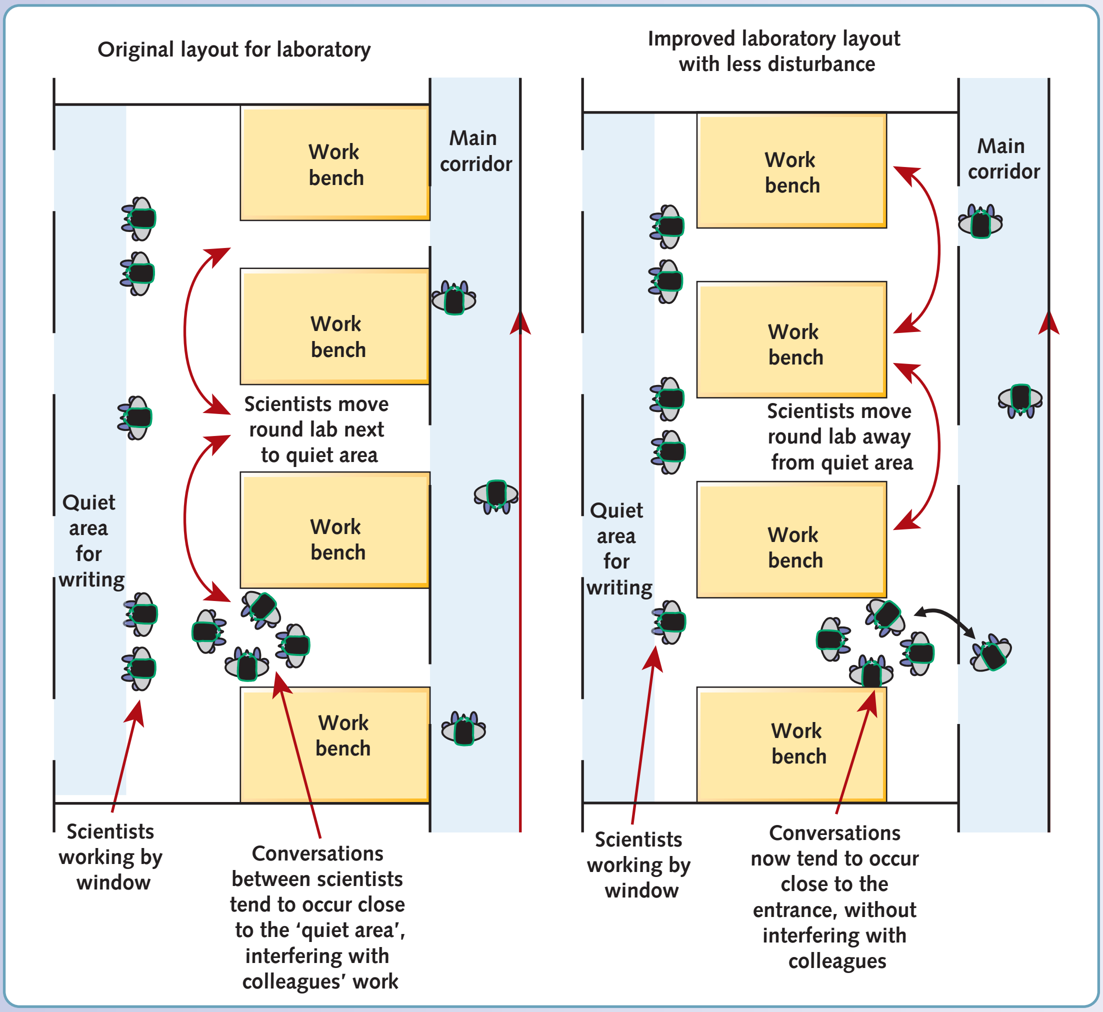


Figure 5.7 Example of an improved laboratory layout that reduces the degree of interference between different types of work (conversations and writing)

(for example, those in advertising agencies or science laboratories) affect behaviour, points out how designing laboratories with busy circulation spaces allows scientists from different research groups to effectively share ideas. People walking around can stop and join a conversation in the doorway of a laboratory. Conversations inside the laboratory, when they are next to where the relatively high-flow movement along the corridor occurs, lead to discussions between research groups.

Figure 5.7 illustrates how laboratory design can, to some extent, reduce the conflict between the benefits of interaction and the need for quiet. The conventional layout on the left allows potentially disruptive conversations to interfere with quiet areas. The marginally modified layout on the right encourages conversations to happen closer to the entrance, without interfering with colleagues.

Functional layout

Functional layout is so called because the functional needs and convenience of the transforming resources that constitute the processes dominate the layout decision. In functional layout, similar activities or resources (or those with similar needs) are located together. This may be because it is convenient to group them together, or that their utilisation can be improved that way. It means that when materials, information or customers flow through the operation, they will take a route from activity to activity according to their needs. Usually, this makes the flow pattern in the operation complex. Examples of process layouts include:

- Hospitals – some processes (e.g. radiography equipment and laboratories) are required by several types of patient.
- Machining the parts for aircraft engines – some processes (e.g. heat treatment) need specialist support (heat and fume extraction); some processes (e.g. machining centres) require the same technical support from specialists; some processes (e.g. grinding machines) get high machine utilisation as all parts that need grinding pass through a single grinding section.
- Supermarkets – some products, such as tinned goods, are convenient to restock if grouped together. Some areas, such as those holding frozen vegetables, need the common technology of freezer cabinets. Others, such as the areas holding fresh vegetables, might be together because that way they can be made to look attractive to customers.

Cell layout

A cell layout is one where materials, information or customers entering the operation are preselected (or preselect themselves) to move to one part of the operation (or cell) in which all the transforming resources, to meet their immediate processing needs, are located. Internally, the cell itself may be arranged in any appropriate manner. After being processed in the cell, the transformed resources may go on to another cell. In effect, cell layout is an attempt to bring some order to the complexity of flow that characterises functional layout. Examples of cell layouts include:

- Some computer component manufacture – the processing and assembly of some types of computer parts may need a special area dedicated to producing parts for one particular customer who has special requirements, such as particularly high quality levels.
- 'Lunch' products area in a supermarket – some customers use the supermarket just to purchase sandwiches, savoury snacks, cool drinks, etc., for their lunch. These products are often located close together in a 'cell' for the convenience of these customers.
- Maternity unit in a hospital – customers needing maternity attention are a well-defined group who can be treated together and who are unlikely to need the other facilities of the hospital at the same time that they need the maternity unit.

Product layout

Product layout involves locating people and equipment entirely for the convenience of the transformed resources. Each product, piece of information or customer follows a prearranged route in which the sequence of required activities corresponds to the sequence in which the processes have been located. The transformed resources 'flow' along a 'line'. This is why this type of layout is sometimes called flow or line layout. Flow is clear, predictable and therefore relatively easy to control. It is the high volume and standardised requirements of the product or service that allow product layouts. Examples of product layout include:

- Car assembly – almost all variants of the same model require the same sequence of processes.
- Self-service cafeteria – generally the sequence of customer requirements (starter, main course, dessert, drink) is common to all customers, but layout also helps control customer flow.

Layout selection

Getting the process layout right is important, if only because of the cost, difficulty and disruption of making any layout change. It is not an activity many businesses would want to repeat very often. Also, an inappropriate layout could mean that extra cost is incurred every time an item is processed. But more than this, an effective layout gives clarity and transparency to the flow of items through a process. There is no better way of emphasising that everyone's activities are really part of an overall process than by making the flow between activities evident to all participants in the process.

One of the main influences on which type of layout will be appropriate is the nature of the process itself, as summarised in its 'process type'. There is often some confusion between process types and layout types, but they are not the same thing. Process types were described earlier in the chapter and indicate a broad approach to the organisation and operation of a process. Layout is a narrower concept but is very clearly linked to process type. Just as process type is governed by volume and variety, so is layout. But for any given process type there are usually at least two alternative layouts. Table 5.1 summarises the alternative layouts for particular process types. Which of these is selected, or whether some hybrid layout is chosen, depends on the relative importance of the performance objectives of the process, especially cost and flexibility. Table 5.2 summarises the advantages and disadvantages.

Table 5.1 Alternative layout types for each process type

<table>
<tr><th>Manufacturing process type</th><th colspan="2">Potential layout types</th><th>Service process type</th></tr>
<tr><td>Project</td><td>Fixed-position layout
Functional layout</td><td>Fixed-position layout
Functional layout
Cell layout</td><td>Professional service</td></tr>
<tr><td>Jobbing</td><td>Functional layout
Cell layout</td><td rowspan="2">Functional layout
Cell layout</td><td rowspan="2">Service shop</td></tr>
<tr><td>Batch</td><td>Functional layout
Cell layout</td></tr>
<tr><td>Mass</td><td>Cell layout
Product layout</td><td rowspan="2">Cell layout
Product layout</td><td rowspan="2">Mass service</td></tr>
<tr><td>Continuous</td><td>Product layout</td></tr>
</table>

Table 5.2 The advantages and disadvantages of the basic layout types

	Advantages	*Disadvantages*
Fixed-position	• Very high mix and product flexibility • Product or customer not moved or disturbed • High variety of tasks for staff	• High unit costs • Scheduling of space and activities can be difficult • Can mean much movement of plant and staff
Functional	• High mix and product flexibility • Relatively robust in the case of disruptions • Relatively easy supervision of equipment or plant	• Low facilities utilisation • Can have very high work-in-progress or customer queueing • Complex flow can be difficult to control
Cell	• Can give a good compromise between cost and flexibility for relatively high-variety operations • Fast throughput • Group work can result in good motivation	• Can be costly to rearrange existing layout • Can need more plant and equipment • Can give lower plant utilisation
Product	• Low unit costs for high volume • Gives opportunities for specialisation of equipment • Materials or customer movement are convenient	• Can have low mix flexibility • Not very robust if there is disruption • Work can be very repetitive

Layout and 'servicescapes'

In high-visibility processes (where the customer 'sees' much of the value-adding processing, see Chapter 1) the general look and feel of the process could be as important, if not more important, than more 'objective' criteria such as cost or distance moved. The term that is often used to describe the look and feel of the environment within an operation or process is its 'servicescape'. There are many academic studies that have shown that the servicescape of an operation plays an important role, both positive and negative, in shaping customers' views of how a high-visibility process is judged. The general idea is that ambient conditions, space factors, and signs and symbols in a service operation will create an 'environmental experience' for both employees and customers, and this environmental experience should support the service concept. The individual factors that influence this experience will then lead to certain responses (in both employees and customers). These responses can be put into three main categories:

- cognitive (what people think);
- emotional (what they feel); and
- physiological (their body's experience).

> OPERATIONS PRINCIPLE
> *Layout should include consideration of the look and feel of the operation to customers and/or staff.*

However, remember that a servicescape will contain not only objective, measurable and controllable stimuli, but also subjective, immeasurable and often uncontrollable stimuli, which will influence customer behaviour. The obvious example is other customers frequenting an operation. As well as controllable stimuli such as the colour, lighting, design, space and music, the number, demographics and appearance of one's fellow customers will also shape the impression of the operation.

Case example

Ducati factory or Google office, they have to look good[7]

Don't be tempted to think that the principles of design that govern the layout and look of an engineering factory and the head office of a global high-tech company should be very different. Where facilities are located within the operation and what they look like is equally important for both. Here are two examples.

The Ducati factory in Bologna

Motorcycles have been built by Ducati at their current factory in Bologna continuously since 1946. But now, as well as producing 350 bikes every day in the busy season, the factory hosts visitors keen to see how the famous motorcycles are made. The assembly process starts in the 'supermarket' area of the factory. Each of the models made by Ducati has its own 'supermarket' zone. Here manufacturing trays are put together with exactly the right parts that will be needed for each of the subsequent stages in the assembly process. Because these 'kits' of parts exactly match the requirements for each product, any parts left in the tray at the end of assembly is an indication that something has gone wrong. Livio Lodi has worked in the Ducati Factory for more than 26 years, and is the official historian and curator of the Ducati museum on the site. He explains: '*With this new style of just-in-time production, we have been able to reduce over 85 per cent of the defects in the final product. . . . Porsche engineers came to us here and explained the way to set up the production in a just-in-time philosophy which they got from Toyota originally.*'

The factory has also become a tourist attraction thanks to the regular and popular factory tours. It is partly a customer relations and marketing device, as well as a production plant. So, the facilities have to be designed to accommodate visitors as well as motorcycle parts. Reviewers of the tour enthuse about the '*fantastic factory tour with wonderful friendly guides*', and how they are '*very informative and knowledgeable*', and '*interesting to see all aspects of the assembly of the wonderful Ducati motorcycle*'. Also, how '*the museum was great and historic with some beautiful old bikes and some up-to-date winners*'. Many visitors are Ducati owners, but the factory and the museum are open to everybody. They say that they don't care which kind of motorcycle you have. But, if you ride a Ducati here, you can park your bike inside the gates. If you don't ride a Ducati, you must leave it outside the entrance.

Google's revolutionary offices

Operations, and therefore operations layouts, are not confined to factories, warehouses, shops and other such workspaces. Financial services, governments, call centres and creative industries all work, for the most part, sitting at desks. (One estimate is that over 70 per cent of the UK's GDP is generated by people working at their desks, though it is admittedly difficult to check.) The layout of offices can affect operations performance for these industries just as much as it can in a factory. For example, Google, like many high-tech companies, is paying much more attention to its employees' work environment, the better to promote creativity and productivity. Google thrives on creativity and it believes that the designs of its offices will provide every employee with a space that will encourage creativity. Google put a lot of time and money into designing what it believes is the perfect work environment: one that can mix business with pleasure in the sense that the staff can relax and unwind during their breaks. The layouts of Google's offices are designed to promote creativity and collaboration. How people move about the space and who they meet and talk to are vital pieces of information that should contribute to any design. The information needs of the process's underlying activities are clearly an important driver of where the various departments of an organisation are located. However, people sometimes are not fully aware how they are interacting with one another, or with the space where they work. So, in addition to examining the formal needs of people's jobs, it is valuable to examine employee behaviour. For example, where do people actually spend the majority of their time? Where and when do the most productive meetings happen? Where and when do people make phone calls? When is the office at its emptiest? When is it most full (and noisy)?

Google was concerned about what the ingredients of its offices should be and how they would all fit together cohesively for a consistent employee experience. Its view was that office design was never just about space, but also about culture, etiquette and rituals. There's a rule at Google that nobody should be located more than 100 metres away from food. In addition to the cafeteria, eco-friendly kitchens, complete with healthy food, are sited around the buildings. There are quiet places, such as libraries and sometimes aquariums, if you want somewhere quiet to relax or think through a problem. Designing these features in office buildings is partly a consequence of the long hours worked by many people in the high-technology industries. Offices must be equipped with areas for working and areas for relaxing (even if that means playing football – an approach championed by Google).

5.4 Diagnostic question: *Are process technologies appropriate?*

Process technologies are the machines, equipment and devices that help processes 'transform' materials, information and customers. It is a particularly important issue because few operations have been unaffected by the advances in process technology over the last two decades. And the pace of technological development is not slowing down. But it is important to distinguish between process technology (the machines and devices that help to create products and services) and product technology (the technology that is embedded within the product or service and creates its specification or functionality). Some process technology, although not used for the actual creation of goods and services, nonetheless plays a key role in facilitating their creation. For example, the information technology systems that run planning and control activities can be used to help managers and operators run the processes. Sometimes this type of technology is called indirect process technology, and it is becoming increasingly important. Many businesses spend more on the computer systems that run their processes than they do on the direct process technology that creates their products and services.

Process technology should reflect volume and variety

Again, different process technologies will be appropriate for different parts of the volume–variety continuum. High-variety, low-volume processes generally require process technology that is general purpose, because it can perform the wide range of processing activities that high variety demands. High-volume, low-variety processes can use technology that is more dedicated to its narrower range of processing requirements. Within the spectrum, from general-purpose to dedicated process technologies, three dimensions in particular tend to vary with volume and variety. The first is the extent to which the process technology carries out activities or makes decisions for itself; that is, its degree of 'automation'. The second is the capacity of the technology to process work; that is, its 'scale' or 'scalability'. The third is the extent to which it is integrated with other technologies; that is, its degree of 'coupling' or 'connectivity'.

> OPERATIONS PRINCIPLE
> ***Process technology in high-volume, low-variety processes is relatively automated, large-scale and closely coupled when compared to that in low-volume, high-variety processes.***

The degree of automation of the technology

To some extent, all technology needs human intervention. It may be minimal – for example the periodic maintenance interventions in a petrochemical refinery. Conversely, the person who operates the technology may be the entire 'brains' of the process, such as a surgeon using keyhole surgery techniques. Generally, processes that have high variety and low volume will employ process technology with lower degrees of automation than those with higher volume and lower variety. For example, investment banks often put together sophisticated financial deals, customised to the needs of individual clients. Each deal may be worth tens or hundreds of millions of dollars. The back office of the bank has to process these deals to make sure that they happen as planned. Much of this processing could be done using relatively general-purpose technology such as spreadsheets. Skilled back-office staff make the decisions rather than the technology. Contrast this with higher-volume, low-variety products, such as straightforward equity (stock) trades. Most of these products are simple and straightforward and are processed in milliseconds, and in very high volume of tens of thousands per day, by 'automated' technology.

The scale/scalability of the technology

There is usually some discretion as to the scale of individual units of technology. For example, the duplicating department of a large office complex may decide to invest in a single, very large, fast printer, or alternatively in several smaller, slower printer/copiers distributed around the operation's various processes. An airline may purchase one or two very wide-bodied aircraft or a larger number of smaller aircraft. The advantage of large-scale technologies is that they can usually process items cheaper than small-scale technologies, but usually need high volume and can cope only with low variety. By contrast, the virtues of smaller-scale technology are often the nimbleness and flexibility that are suited to high-variety, lower-volume processing. For example, four small machines can between them produce four different products simultaneously (albeit slowly), whereas a single large machine with four times the output can produce only one product at a time (albeit faster). Small-scale technologies can also be more robust. Suppose the choice is between three small machines and one larger one. In the first case, if one machine breaks down only a third of the capacity is lost, but in the second, capacity is reduced to zero.

The equivalent to scale for some types of information processing technology is scalability. By scalability we mean the ability to shift to a different level of useful capacity quickly, and cost-effectively. Scalability is similar to absolute scale in so much as it is influenced by the same volume–variety characteristics. IT scalability relies on consistent IT platform architecture and the high process standardisation that is usually associated with high-volume and low-variety operations.

The coupling/connectivity of the technology

Coupling means the linking together of separate activities within a single piece of process technology to form an interconnected processing system. Tight coupling usually gives fast process throughput. For example, in an automated manufacturing system, products flow quickly without delays between stages, and inventory will be lower – it can't accumulate when there are no 'gaps' between activities. Tight coupling also means that flow is simple and predictable, making it easier to keep track of parts when they pass through fewer stages, or information when it is automatically distributed to all parts of an information network. However, closely coupled technology can be both expensive (each connection may require capital costs) and vulnerable (a failure in one part of an interconnected system can affect the whole system). The fully integrated manufacturing system constrains parts to flow in a predetermined manner, making it difficult to accommodate products with very different processing requirements. So, coupling is generally more suited to relatively low variety and high volume. Higher-variety processing generally requires a more open and unconstrained level of coupling because different products and services will require a wider range of processing activities.

New technology is changing the diagonal

Figure 5.8 illustrates these three dimensions of process technology and the implied 'diagonal' between low volume–high variety, and high volume–low variety. It also shows how some developments in technology have, to some extent, overcome the implied trade-off between flexibility and cost. Specifically, digitisation and the vastly increased computing power embedded in many new technologies has made it easier to achieve lower costs while not having to sacrifice the ability to provide variety. For example, the IT systems, databases and algorithms behind internet banking services allow customers to access a very wide variety of customised services, while retaining or enhancing back-office efficiency. Similarly, so-called 3D knitting machines can produce an entire garment including the arms, collars and other parts that would normally be produced separately and sewn on, using a single machine with a single thread.

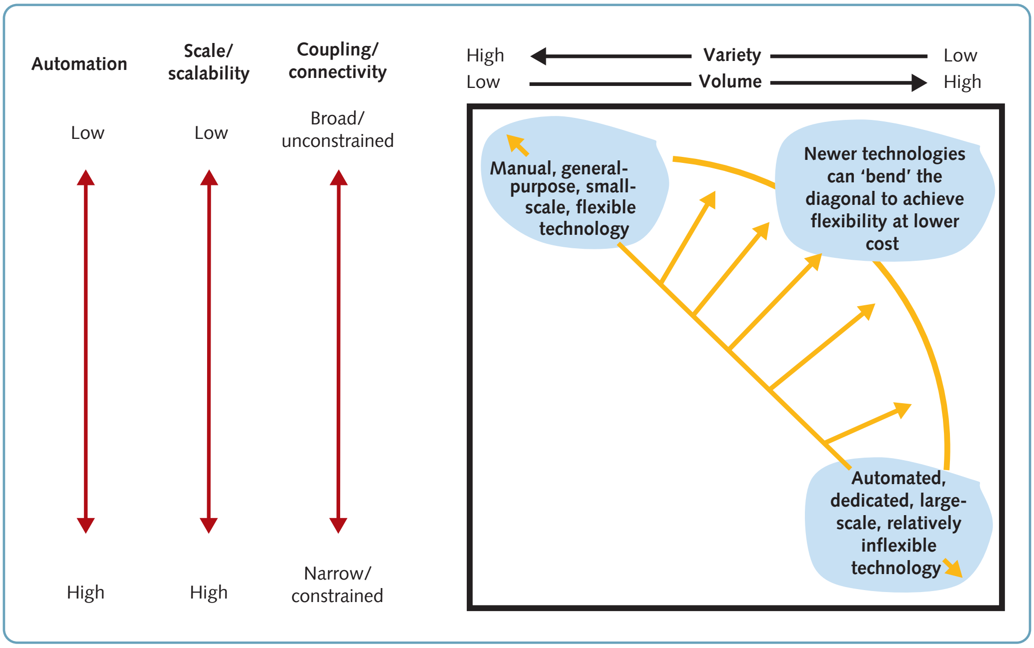


Figure 5.8 Different process technologies are important for different volume–variety combinations but some newer technologies can achieve both flexibility and low cost

Not only does this allow a wide variety of garments to be produced at relatively low cost, it reduces the waste of material. The effect of this is to 'bend' what has hitherto been accepted as the 'natural' diagonal.

5.5 Diagnostic question: *Are job designs appropriate?*

Job design is about how people carry out their tasks within a process. It defines the way they go about their working lives, it positions the expectations of what is required of them, and it influences their perceptions of how they contribute to the organisation. It also defines their activities in relation to their work colleagues and it channels the flow of communication between different parts of the operation. But of most importance, it helps to develop the culture of the organisation: its shared values, beliefs and assumptions. Inappropriately designed jobs can destroy the potential of a process to fulfil its objectives, no matter how appropriate its layout or process technology. So jobs must be designed to fit the nature of the process. However, before considering this, it is important to accept that some aspects of job design are common to all processes, irrespective of what they do or how they do it. Consider the following:

- *Safety.* The primary and universal objective of job design is to ensure that all staff performing any task within a process are protected against the possibility of physical or mental harm.
- *Ethical issues.* No individual should be asked to perform any task that is either illegal or (within limits) conflicts with strongly held ethical beliefs.
- *Work–life balance.* All jobs should be structured so as to promote a healthy balance between time spent at work and time away from work.

Note that all these objectives of job design are also likely to improve overall process performance. However, the imperative to follow such objectives for their own sake transcends conventional criteria.

Job design should reflect volume and variety

As with other aspects of process design, the nature and challenges of job design are governed largely by the volume–variety characteristics of a process. An architect designing major construction projects will perform a wide range of very different, often creative and complex tasks, many of which are not defined at the start of the process, and most of which have the potential to give the architect significant job satisfaction. By contrast, someone in the architect's accounts office keying in invoice details has a job that is repetitive, has little variation, is tightly defined and is one that cannot rely on the intrinsic interest of the task itself to maintain job commitment. These two jobs will have different characteristics because they are parts of processes with different volume and variety positions. Three aspects of job design in particular are affected by the volume–variety characteristics of a process: how tasks are to be allocated to each person in the process; the degree of job definition; and the methods used to maintain job commitment. Figure 5.9 illustrates this.

> OPERATIONS PRINCIPLE
> *Job designs in high-volume, low-variety processes are relatively closely defined with little decision-making discretion and needing action to encourage commitment when compared to those in low-volume, high-variety processes.*

The division of labour – how should tasks be allocated?

The most obvious aspect of any individual's job is how big it is; that is, how many of the tasks within any process are allocated to an individual. Should a single individual perform all the process? Alternatively, should separate individuals or teams perform each task? Separating tasks into smaller parts between individuals is called the division of labour. Perhaps its epitome is the assembly line, where products move along a single path and are built up by operators continually repeating a single task. This is the predominant model of job design in

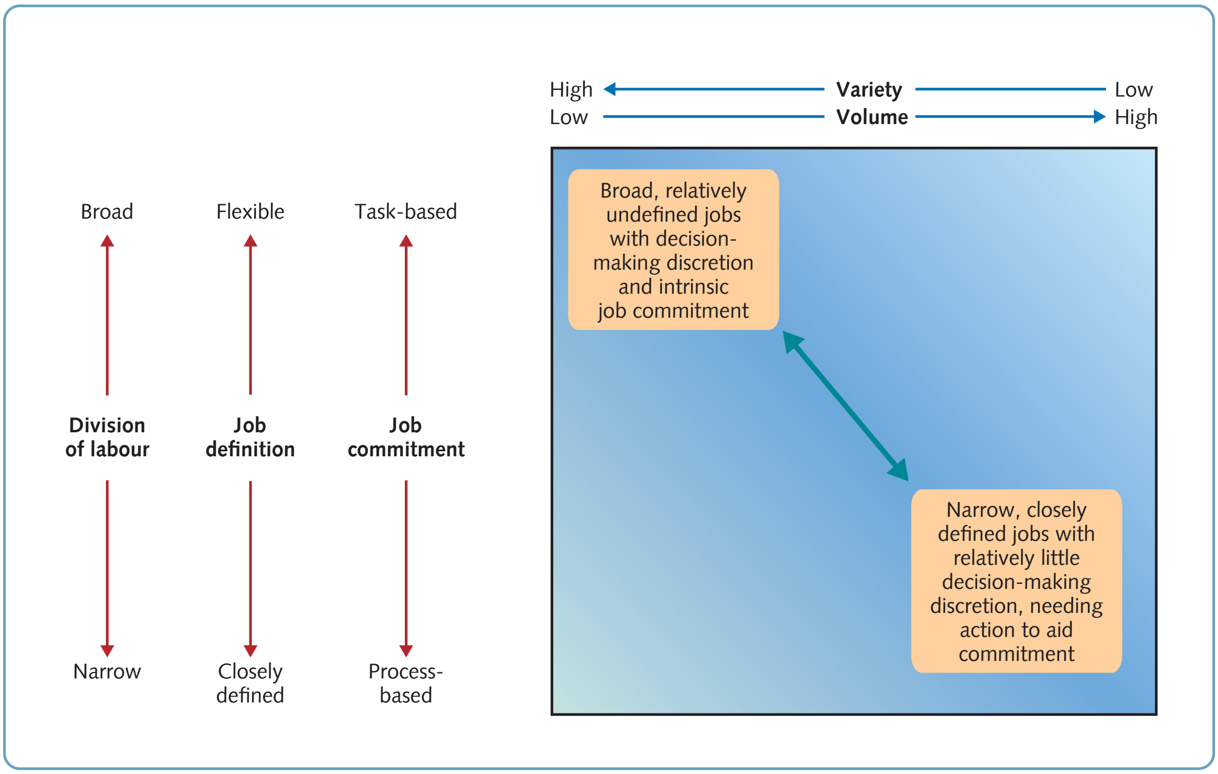


Figure 5.9 Different job designs are appropriate for different volume–variety combinations

most high-volume, low-variety processes. For such a process there are some real advantages in division-of-labour principles:

- *It promotes faster learning*. It is obviously easier to learn how to do a relatively short and simple task than a long and complex one, so new members of staff can be quickly trained and assigned to their tasks.
- *Automation becomes easier*. Substituting technology for labour is considerably easier for short and simple tasks than for long and complex ones.
- *Non-productive work is reduced*. In large, complex tasks the proportion of time between individual value-adding elements can be very high; for example, in manufacturing, picking up tools and materials, putting them down again and generally searching and positioning.

There are also serious drawbacks to a highly divided job:

- *It is monotonous*. Repeating the same task, eight hours a day and five days a week, is not fulfilling. This may lead to an increased likelihood of absenteeism, staff turnover and error rates.
- *It can cause physical injury*. The continued repetition of a very narrow range of movements can, in extreme cases, lead to physical injury. The over-use of some parts of the body (especially the arms, hands and wrists) can result in pain and a reduction in physical capability, called repetitive strain injury (RSI).
- *It can mean low flexibility*. Dividing up a task into many small parts often gives the job design a rigidity that is difficult to adapt under changing circumstances. For example, if an assembly line has been designed to make one particular product but then has to change to manufacture a quite different product, the whole line will need redesigning. This will probably involve changing every operator's set of tasks.
- *It can mean poor robustness*. Highly divided jobs imply items passing between several stages. If one of these stages is not working correctly, for example because some equipment is faulty, the whole operation is affected. On the other hand, if each person is performing the whole of the job, any problems will affect only that one person's output.

To what degree should jobs be defined?

Jobs in high-variety processes are difficult to define in anything but the most general terms. Such jobs may require tacit knowledge gained over time and through experience, and often require individuals to exercise significant discretion in what they do and how they do it. Some degree of job definition is usually possible and advisable, but it may be stated in terms of the 'outcome' of the task rather than in terms of the activities within the task. For example, the architect's job may be defined in terms of 'achieving overall coordination, taking responsibility for articulating the overall vision of the project, ensuring stakeholders are comfortable with the process, etc'. By contrast, a process with less variety and higher volume is likely to be defined more closely, with the exact nature of each activity defined and individual staff trained to follow a job step by step.

How should job commitment be encouraged?

Many factors may influence job commitment. An individual's job history and expectations, relationships with co-workers, personal circumstances, can all be important. So too are the volume and variety characteristics of the process, by defining the possible ways in which commitment can be enhanced. In high-variety processes, especially those with a high degree of staff discretion, job commitment is likely to come from the intrinsic nature of the task itself. Exercising skill and decision-making, for example, can bring its own satisfaction. Of course, commitment can be enhanced through extra responsibility, flexibility in working times and so on, but the main motivator is the job itself. By contrast, low-variety, high-volume jobs, especially those designed with a high division of labour and little discretion, can be highly alienating. Such jobs have relatively little intrinsic task satisfaction. It has to be 'designed into' the process by emphasising the satisfaction to be gained from the performance of the process overall. A number of job design approaches have been suggested for achieving this in processes involving relatively repetitive work.

Job enlargement

Job enlargement involves allocating a larger number of tasks to individuals, usually by combining tasks that are broadly of the same type as those in the original job. This may not involve more demanding or fulfilling tasks, but it may provide a more complete and therefore slightly more meaningful job. If nothing else, people performing an enlarged job will not repeat themselves as often. For example, suppose that the manufacture of a product has traditionally been split up on an assembly line basis into ten equal and sequential jobs. If that job is then redesigned so as to form two parallel assembly lines of five people, each operator would have twice the number of tasks to perform.

Job enrichment

Like job enlargement, job enrichment increases the number of tasks in a job, but also implies allocating tasks that involve more decision-making, or greater autonomy, and therefore greater control over the job. This could include the maintenance of, and adjustments to, any process technology used, the planning and control of activities within the job, or the monitoring of quality levels. The effect is both to reduce repetition in the job and to increase personal development opportunities. So, in the assembly line example, each operator working on a job could also be allocated responsibility for carrying out routine maintenance and such tasks as record-keeping and managing the supply of materials.

Job rotation

Job rotation means moving individuals periodically between different sets of tasks to provide some variety in their activities. When successful, job rotation can increase skill flexibility and make a small contribution to reducing monotony. However, it is not always viewed as beneficial either by management (because it can disrupt the smooth flow of work) or by the people performing the jobs (because it can interfere with their rhythm of work).

Empowerment

Empowerment means enhancing individuals' ability, and sometimes authority, to change how they do their jobs. Some technologically constrained processes, such as those in chemical plants, may limit the extent that staff can dilute their highly standardised task methods without consultation. Other less-defined processes to empowerment may go much further.

Team working

Closely linked to empowerment, team-based work organisation (sometimes called self-managed work teams) is where staff, often with overlapping skills, collectively perform a defined task and have some discretion over how they perform the task. The team may control such things as task allocation between members, scheduling work, quality measurement and improvement, and sometimes even the hiring of staff. Groups are described as 'teams' when the virtues of working together are being emphasised and a shared set of objectives and responsibilities is assumed.

Work–life balance

A number of factors have made it increasingly difficult to separate work life from personal life: the decline in the number of people with clearly delineated working times; less distinction between what are clearly 'work' technologies (laptops, mobile devices, etc.) and personal devices; and an increase in home working. All of these make maintaining an appropriate split between work and personal life more difficult. But reaching a healthy work–life balance is important because, among other benefits, employee retention is improved, an operation's reputation in the jobs market is enhanced (thus helping to attract more able staff), staff physical or mental health problems are reduced and, arguably, staff effectiveness is increased.

Critical commentary

There could be three sets of criticisms prompted by the material covered in this chapter. The first relates to the separation of process design into two parts – positioning and analysis. It can reasonably be argued that this separation is artificial in so much as (as is admitted at the beginning of this chapter) the two approaches are very much interrelated. An alternative way of thinking about the topic would be to consider all aspects of the arrangement of resources together. This would include the issues of layout that have been discussed in this chapter, together with the more detailed process-mapping issues described in Chapter 6. The second criticism would challenge the core assumption of the chapter – that many significant process design decisions are influenced primarily by volume and variety. Whereas it is conventional to relate layout and (to a slightly lesser extent) process technology to volume–variety positioning, it is less conventional to do so for issues of job design. Some would argue that the vast majority of job design decisions will not vary significantly with volume and variety. The final criticism is also related to job design. Some academics would argue that our treatment of job design is over-influenced by the discredited (in their eyes) principles of the 'scientific' management movement that grew into 'work study' and 'time and motion' management.

SUMMARY CHECKLIST

- ☐ Do processes match volume–variety requirements?
- ☐ Are 'process types' understood and do they match volume–variety requirements?
- ☐ Can processes be positioned on the 'diagonal' of the product–process matrix?
- ☐ Are the consequences of moving away from the 'diagonal' of the product–process matrix understood?
- ☐ Are the implications of choosing an appropriate layout, especially the balance between process flexibility and low processing costs, understood?
- ☐ Are the process layouts appropriate?
- ☐ Which of the four basic layout types that correspond to different positions on the volume–variety spectrum is appropriate for each process?
- ☐ Have 'soft' servicescape factors been considered?
- ☐ Is process technology appropriate?
- ☐ Is the effect of the three dimensions of process technology (the degree of automation, the scale/scalability, and the coupling/connectivity of the technology) understood?
- ☐ Are job designs appropriate?
- ☐ Does job design ensure the imperative to design jobs that are safe, ethical and promote an adequate work/life balance?
- ☐ Is the extent of division of labour in each process appropriate for its volume–variety characteristics?
- ☐ Is the extent of job definition in each process appropriate for its volume–variety characteristics?
- ☐ Are job commitment mechanisms in each process appropriate for its volume–variety characteristics?

Case study

McPherson Charles Solicitors

Grace Whelan, Managing Partner of McPherson Charles, welcomed the three solicitors into the meeting room. She outlined the agenda, essentially their thoughts and input into the rolling three-year plan. McPherson Charles, based in Bristol in the West of England, had grown rapidly to be one of the biggest law firms in the region, with 21 partners and around 400 staff. It was an ambitious partnership aiming to maintain its impressive growth record. The firm was managed through 15 teams, each headed by a partner. The meeting was intended to be the first stage in 'Plans for the Future', a programme to improve the effectiveness of the firm's operations. The three partners attending the meeting with Grace were Simon Reece (Family Law), Kate Hutchinson (Property) and Hazel Lewis (Litigation). Grace asked for ideas on what the firm should prioritise in order to improve its performance further.

Simon Reece kicked things off: '*I think the first thing we need to agree on is that, for a professional service firm like ourselves, the quality of our people will always be the most important issue. We need to be absolutely confident that our staff not only have the best possible understanding of their own branch of the law, but also have the necessary client-relationship skills to consolidate our business position with increasingly demanding clients.*' Hazel Lewis was not so sure: '*Of course I agree that the quality of our staff is an important issue, but that has always been true. What is new is the help we can get from some serious investment in technology and software. Just getting our systems and processes right would, I am sure, save us a lot of time and effort, and of course, reduce our cost base*.'

'*I really don't think spending more money on technology is the answer Hazel*', Simon continued. '*We need more time to really understand our clients and being process and IT focused just doesn't work for us, we need another way of managing. The key is increasing revenue, not penny pinching about costs, and to do that we need to really concentrate on relationship skills. Family law is like walking through a minefield; you can easily offend clients who are, almost by definition, in a highly emotional state. I think we need to make sure that senior members of staff with experience of managing client relationships pass on their knowledge to those who are less experienced.*'

'*I disagree, Simon.*' It was Kate Hutchinson now. '*Our clients really are increasingly cost conscious and if we don't deliver value for money, word will spread very fast and our business will dry up. Much of the time we over-engineer our services. Why should we use highly qualified and expensive lawyers for every single task? I am convinced that, with slick systems and enhanced training, non-qualified people could do much of our work.*'

Grace knew that solicitors liked nothing better than disagreeing – it was what they did best – and she knew that this was going to be another long meeting.

In very simple terms, these are the types of activities that each team was engaged in.

Simon Reece, Family Law

'*We are called the "family law" team but basically what we do is to help people through the trauma of divorce, separation and break up. Our biggest "high value" clients come to us because of word-of-mouth recommendation. Last year we had 89 of these "high value" clients and they all valued the personal touch that we were able to give them, getting to know them well and spending time with them to understand the, often "hidden" aspects of their case. These interviews cannot be rushed. These clients tend to be wealthy people and we will often have to drop everything and go off half-way round the world to meet and discuss their situation. There are no standard procedures, every client is different, and everyone has to be treated as an individual. So we have a team of individuals who rise to the challenge each time and give great service. Of course, not all our clients are the super-rich. About a third of our annual family law income comes from about 750 relatively routine divorce and counselling cases. This work is a lot less interesting and I try to make sure that all my team have a mix of interesting and routine work over the year. I encourage them to exercise and develop their professional judgement. They are empowered to deal with any issues themselves or call in myself or one of the more senior members of the team for advice if appropriate. It is important to give this kind of responsibility to them so that they see themselves as part of a team. We are also the only part of the firm that has adopted an open-plan office arrangement centred around our specialist library of family case law.*'

Hazel Lewis, Litigation

'*The litigation team provides a key service for our commercial client base. Our primary work consists of handling bulk collections of debt. The group has 17 clients, of which 5 provide 85 per cent of the total volume. We work closely with the accounts departments of the client companies and have developed a semi-automatic approach to debt collection. Staff input data received from their clients into the system; from that point everything progresses through a predefined process, letters are produced, queries responded to and eventually debts collected, ultimately through court proceedings if necessary. Work tends to come in batches from clients and varies according to economic conditions, time of year and client sales activities. At the moment things are fairly steady; we had 872 new cases last week. The details of each case are sent over by the client; our people input the data onto our screens and set up a standard diary system for sending letters out. Some people respond quickly to the first letter and often* *the case is closed within a week or so; other people ignore letters and eventually we initiate court proceedings. We know exactly what is required for court dealings and have a pretty good process to make sure all the right documentation is available on the day.'*

Kate Hutchinson, Property

'We are really growing fast and are building up an excellent reputation locally for being fast, friendly and giving value for money. Most of our work is "domestic", acting for individuals buying or selling their home, or their second home. Each client is allocated to a solicitor who becomes their main point of contact. But, given that we can have up to 100 domestic clients a week, most of the work is carried out by the rest of the team behind the scenes. There is a relatively standard process to domestic property sales and purchases and we think that now we are pretty efficient at managing these standard jobs. Our process has four stages: one dealing with land registry searches, one liaising with the banks that are providing the mortgage finance, one that makes sure surveys are completed and one section that finalises the whole process to completion. We believe that this degree of specialisation can help us achieve the efficiencies that are becoming important as the market gets more competitive. Increasingly, we are also getting more complex "special" jobs. These are things like "volume re-mortgage" arrangements; rather complex "one-off" jobs, where a mortgage lender transfers a complex set of loan assets to another lender. "Special" jobs are always more complex than the domestic work and sometimes there are times when fast completion is particularly important and that can throw us a bit. The firm has recently formed partnerships with two large speculative builders, so we are getting into special "plot sales". All these "specials" do involve a lot of work and can occupy several members of the team for a time. We are now getting up to 25 of these "specials" each week, and they can be somewhat different to our normal work but we try to follow roughly the same process with them as the normal domestic jobs.'

Thomas Barwick/Stone/Getty Images

Are each team's processes appropriate?

Grace was concerned. The three teams obviously had to cope with very different volumes of work and variety of activities. It was also clear that each team had developed different approaches to managing their processes. The question that she needed to address was whether each team's approach was appropriate for the demands placed upon it.

Questions

1. What are the individual 'services' offered by each of the three teams?
2. Where would you place each service in a scale that goes from relatively low volume/relatively high variety, to relatively high volume/relatively low variety?
3. How would you describe each team's process in terms of its layout, the technology (if any) it uses, and the job designs of its staff?
4. Use the above information to draw a 'product–process matrix'. What does it indicate?

Applying the principles

Some of these exercises can be answered by reading the chapter. Others will require some general knowledge of business activity and some might require an element of investigation. **Model answers for the first two questions are available on the student companion website.** *For model answers to the other questions in this section, please ask your tutor/lecturer.*

1. Visit a branch of a retail bank and consider the following questions:

a) What categories of service does the bank seem to offer?

b) To what extent does the bank design separate processes for each of its types of service?

c) What are the different process design objectives for each category of service?

2. A company that produces a wide range of specialist educational kits for five–ten-year-olds is based in an industrial unit arranged in a simple layout with six departments, each performing a separate task. The layout is shown in Figure 5.10, together with the results of an investigation of the flow of parts and products between each department. However, the company plans to revamp its product range. '*This new range will totally replace our existing products, and although I believe our existing layout is fine for the current product range, I think that we will need to reconfigure our layout when we make the transition to the new product range.*' (COO of the company) The estimate for the flow between the departments when the new product range is introduced is shown in Figure 5.11.

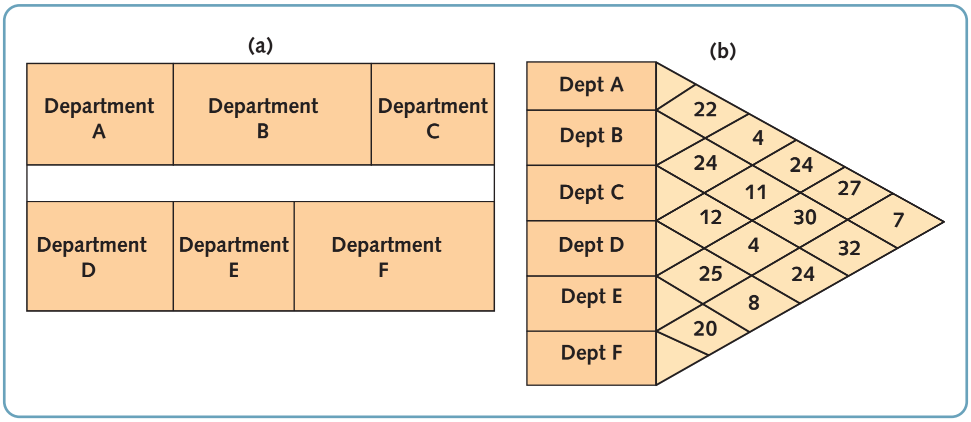


Figure 5.10 (a) The current layout of the educational kits producer, and (b) the current interdepartmental flow of parts and products (in pallet loads)

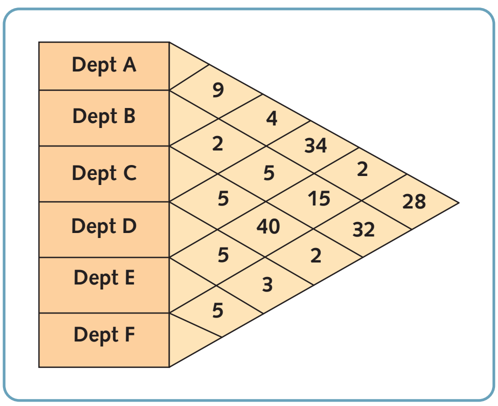


Figure 5.11 Estimated interdepartmental flow when the new product range is introduced

a) Is the COO right in thinking that the current layout is right for the current product range?

b) Assuming that the estimate of future interdepartmental flow is correct, how would you rearrange the factory?

3. The International Frozen Pizza Company (IFPC) operates in three markets globally. Market 1 is its largest market, where it sells 25,000 tons of pizza per year. In this market it trades under the name 'Aunt Bridget's Pizza' and positions itself as making pizza 'just as your Aunt Bridget used to make' (apparently she was good at it). It is also known for innovation, introducing new and seasonal pizza toppings on a regular basis. Typically, it sells around 20 varieties of pizza at any one time. Market 2 is smaller, selling around 20,000 tons per year under its 'Poppet's Pizza' brand. Although less innovative than Market 1, it still sells around 12 varieties of pizza. Market 3 is the smallest of the three, selling 10,000 tons per year of relatively high-quality pizzas under its 'Deluxe Pizza' brand. Like Aunt Bridget's Pizza, Deluxe Pizza also sells a relatively wide product range for the size of its market. Currently, both Markets 1 and 3 use relatively little automation and rely on high numbers of people, employed on a shift-system, to assemble their products. Market 2 has always been keen to adopt more automated production processes and uses a mixture of automated assembly and manual assembly. Now the management in Market 2 has developed an almost fully automated pizza assembly system (APAS). They claim that the APAS could reduce costs significantly and should be adopted by the other markets. Both Markets 1 and 3 are sceptical. 'It may be cheaper, but it can't cope with a high variety of products', is their response.

a) Use the product–process matrix to explain the proposal by the management of Market 2.

4. A direct marketing company sells kitchen equipment through a network of local representatives working from home. Typically, individual orders usually contain 20–50 individual items. Much of the packing process is standardised and automatic. The vice-president of distribution is proud of his distribution centre: '*We have a slick order fulfilment operation with lower costs per order, few packing errors, and fast throughput times. Our main problem is that the operation was designed for high volumes but the direct marketing business using representatives is, in general, on a slow but steady decline.*' Increasingly, customers are moving towards using the company's recently launched website or just buying from supermarkets and discount stores. Bowing to the inevitable, the company has started selling its products through discount stores. The problem is how to distribute its products through these new channels. Should it modify its existing fulfilment operation or subcontract the business to specialist carriers? '*Although our system is great at what it does, it would be difficult to cope with very different types of order. Website orders will mean dealing with a far greater number of individual customers, each of whom will place relatively small orders for one or two items. We are not designed to cope with that kind of order. We would have the opposite problem delivering to discount stores. There, comparatively few customers would place large orders for a relatively narrow range of products. That is the type of job for a conventional distribution company. Another option would be to accept an offer from a non-competitor company who sell their products in a very similar way*.'

a) What are the implications of the different sales channels for the existing distribution centre?

5. Revisit the case example that examines Legal and General's modular housing venture (see p. 169). Does their use of a factory to 'build' houses invalidate the idea that volume and variety govern the nature of operations processes?

6. Visit a supermarket and observe people's behaviour. You may have to exercise some discretion when doing this; people generally don't like to be stalked a round the supermarket too obviously.

a) What layout type is a conventional supermarket and how does it differ from a manufacturing operation using the same layout type?

b) Some supermarkets are using customer-tracking technology that traces the flow of customers through the shop. What are the benefits of using this type of technology for supermarkets?

Notes on chapter

1 The information on which this example is based is taken from: Sutherland, E. (2017) 'Weihai and mighty', Drapersonline, 16 June; and www.sandsfilms.co.uk [accessed 16 September 2020].

2 'Global fashion industry statistics – international apparel', Fashion United, http://www.fashionunited.com/global-fashion-industry-statistics-international-apparel [accessed 16 September 2020].

3 Quoted on Dishang Group's website, www.dishang-group.com [accessed 16 September 2020].

4 The information on which this example is based is taken from: Wilmore, J. (2019) 'We take a look around L&G's housing factory', *Inside Housing*, 14 February, https://www.insidehousing.co.uk/insight/insight/we-take-a-look-around-lgs-housing-factory-60136 [accessed 16 September 2020]; Legal and General website, https://www.legalandgeneral.com/modular/a-modern-method/ [accessed 16 September 2020]

5 'The built environment: Achieving a resilient recovery with the circular economy', report by the Ellen MacArthur Foundation, https://www.ellenmacarthurfoundation.org/our-work/activities/covid-19/policy-and-investment-opportunities/the-built-environment [accessed 4 October 2020].

6 The information on which this example is based is taken from: Booth, R. (2017) 'Francis Crick Institute's £700m building too noisy to concentrate', *The Guardian*, 21 November.

7 The information on which this example is based is taken from: Urry, J. (2017) 'Inside Ducati: MCN walk around the Bologna factory', *Motorcycle News*, 21 September; Blain, L. (2014) 'Inside the Ducati factory building the 1199 Panigale from the ground up', Newatlas.com, 23 January; Segran, E. (2015) 'Designing a happier office on the super cheap', Fast Company, 30 March; Hickey, S. (2014) 'Death of the desk: The architects shaping offices of the future', *The Guardian*, 14 September.

Taking it further

Bock, L. *(2015) Work Rules! Insights from inside Google that will transform how you live and lead, John Murray.* *With an agenda far wider than this chapter, it is nevertheless an absorbing book that gives an insight into an absorbing firm.*

Boxall, P. and Winterton, J. *(2018) 'Which conditions foster high-involvement work processes? A synthesis of the literature and agenda for research', Economic and Industrial Democracy, 39 (1), pp. 27–47.* *Academic, but comprehensive and insightful.*

Brynjolfsson, E. and McAfee, A. *(2014) The Second Machine Age: Work, progress, and prosperity in a time of brilliant technologies, W. W. Norton & Company.* *This is one of the most influential recent books on how technology will change our lives.*

Hopp, W.J. and Spearman, M.L. *(2011) Factory Physics, 3rd edition, Waveland Press.* *Very technical so don't bother with it if you aren't prepared to get into the maths. However, some fascinating analysis, especially concerning Little's Law.*

Jeston, J. and Nelis, J. *(2014) Business Process Management, 3rd edition, Routledge.* *Written by two consultants, it gives chapter and verse on the orthodoxy of business process management – a broader topic than is covered in this chapter, but useful nonetheless.*

Panagacos, T. *(2012) The Ultimate Guide to Business Process Management: Everything you need to know and how to apply it to your organization, CreateSpace Independent Publishing Platform.* *It is very much a practitioner book. It tells you how to do it.*

Shorrock, S. and Williams, C. *(eds) (2016) Human Factors and Ergonomics in Practice: Improving system performance and human well-being in the real world, Routledge.* *An edited book, but with lots of examples of the real practice of human factors and ergonomics.*

Sirgy, M.J. and Lee, D.-J. *(2018) 'Work–life balance: An integrative review', Applied Research in Quality of Life, 13 (1), pp. 229–254.* *Wide-ranging summary of the research in the area.*

6 Process design 2 – Analysis

Introduction

The previous chapter set the broad parameters for process design; in particular it showed how volume and variety shape the resources of the process in terms of its appropriate layout, process technology and the design of its jobs. But this is only the beginning of process design. Within these broad resource parameters there are many more detailed decisions regarding the activities carried out by the process that will dictate the way materials, information and customers flow through the process. Do not dismiss these detailed design decisions as merely the 'technicalities' of process design. They are important because they go a long way to determining the actual performance of the process in practice and eventually its contribution to the performance of the whole business (see Figure 6.1).

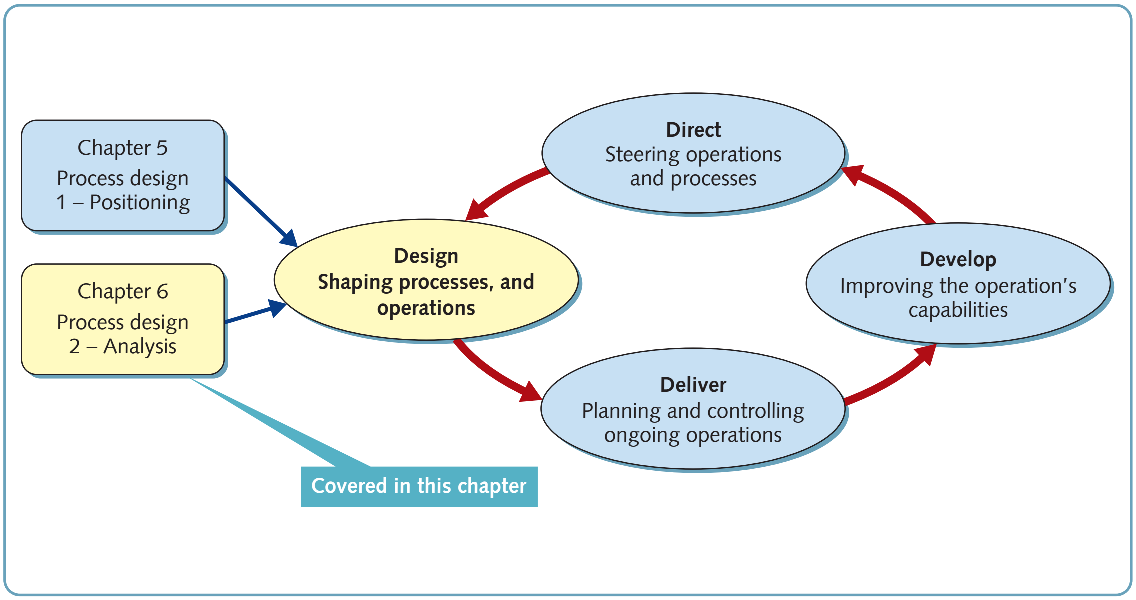


Figure 6.1 **Process design – analysis involves calculating the details of the process, in particular its objectives, sequence of activities, allocation of tasks and capacity, and its ability to incorporate the effects of variability**

EXECUTIVE SUMMARY

6.1 Is the importance of detailed process design understood?

↓

6.2 Are process performance objectives clear?

↓

6.3 How are processes currently designed?

↓

6.4 Are process tasks and capacity configured appropriately?

↓

6.5 Is process variability recognised?

6.1 Is the importance of detailed process design understood?

How individual processes are designed is important to any operation because they are the building blocks of an operation. They link together to form 'internal process chains'. However, no chain is stronger than its weakest link, so the output from an operation will be limited by the effectiveness of its weakest process. This is why process design is so important. It involves calculating the details of the process, in particular its objectives, sequence of activities, allocation of tasks and capacity, and its ability to incorporate the effects of variability. It is the complementary activity to the broad positioning of processes that was described in the previous chapter.

6.2 Are process performance objectives clear?

The major objective of any process in the business is to support the business's overall objectives. Process design must therefore reflect the relative priority of the normal performance objectives of quality, speed, dependability, flexibility and cost. At a more detailed level, process design defines the way units flow through an operation. Therefore, more 'micro' performance objectives are also useful in process design. Four in particular are used. These are throughput (or flow) rate, throughput time, the number of units in the process (work-in-process) and the utilisation of process resources.

6.3 How are processes currently designed?

Much process design is in fact redesign, and a useful starting point is to fully understand how the current process operates. The most effective way of doing this is to map the process in some way. This can be done at different levels using slightly different mapping techniques. Sometimes it is useful to define the degree of visibility for different parts of the process, indicating how much of the process is transparent to customers.

6.4 Are process tasks and capacity configured appropriately?

This is a complex question with several distinct parts. First, it is necessary to understand the task precedence to be incorporated in the process. This defines what activities must occur before others. Second, it is necessary to examine how alternative process design options can incorporate series and parallel configuration. These are sometimes called 'long–thin' and 'short–fat' arrangements. Third, cycle time and process capacity must be calculated. This can help to allocate work evenly among the stages of the process (called balancing). Fourth, the relationship between throughput, cycle time and work-in-process must be established. This is done using a simple but extremely powerful relationship known as Little's Law (throughput time = work-in-process × cycle time).

6.5 Is process variability recognised?

In reality, processes have to cope with variability, both in terms of time and the tasks that are performed within the process. This variability can have very significant effects on process behaviour, usually to reduce process efficiency. Queuing theory can be used to understand this effect. In particular, it is important to understand the relationship between process utilisation and the number of units waiting to be processed (or throughput time).

6.1 Diagnostic question: *Is the importance of detailed process design understood?*

To 'design' is to conceive the looks, arrangement and workings of something before it is constructed. Process design should be treated at two levels – the broad, aggregated level and the more detailed level. The previous chapter took a broad approach by relating the resources in the process to the volume–variety position of the process. That will have identified the broad process type, and given some guidance as to the layout, process technology and job designs to be used within the process. This chapter takes a more detailed view of the activities undertaken by the process. However, in working out the details of a process design it is sometimes necessary to revisit the overall broad assumptions under which it is being designed. This is why the detailed analysis of process design covered in this chapter should always be thought through in the context of the broader process positioning issues covered in Chapter 5. The following case example illustrates processes whose detailed design is important in determining their effectiveness.

How individual processes are designed is important to any operation because they are its building blocks. They link together to form 'internal process chains' through which flow material, information or people. However, no chain is stronger than its weakest link. The output from a process is limited by the effectiveness of its worst stage and no operation is better than its weakest process. Configuring internal process networks involves designing each process and integrating individual processes to form an effective network. The root causes of business failure are often the result of the failures of everyday operational processes. And a prerequisite for long-term success is getting process design right.

Case example

Changi Airport[1]

Airports are complex operations. Really complex. Their processes handle passengers, aircraft, crew, baggage, commercial cargo, food, security, restaurants and numerous customer services. The operations managers, who oversee the daily operations of an airport, must cope with aviation administration rules and regulations, a huge number of airport service contracts, usually thousands of staff with a wide variety of specialisms, airlines with sometimes competing claims to service priority, customers, some of whom fly every week, and others who are a family of seven with two baby strollers who fly once a decade. Their processes are also vulnerable to disruptions from late arrivals, aircraft malfunction, weather, the industrial action of workers two continents away, conflicts, terrorism and exploding volcanoes in Iceland. Designing the processes that can operate under these conditions must be one of the most challenging operations tasks. So to win prizes for 'Best Airport' customer service and operating efficiency year after year has to be something of an achievement, which is what the sixth-busiest international airport, Changi Airport in Singapore, has done. As a major air hub in Asia, Changi serves more than 100 international airlines flying to some 300 cities in about 70 countries and territories worldwide. It handles almost 60 million passengers per year (that's roughly 10 times the size of Singapore's population). A flight takes off or lands at Changi roughly once every 90 seconds.

When Changi opened its new Terminal 4, it increased the airport's annual passenger handling capacity to around 82 million. Every stage of the customers' journey through the terminal was designed to be as smooth as possible. The aim of all the processes within and around the terminal was to provide fast, smooth and seamless flow for passengers. Each stage in the customer journey was provided with enough capacity to cope with anticipated demand. A new overhead bridge was built across the airport boulevard connecting T4 with Singapore's highway system to enable the movement of cars, buses and airside vehicles. Two new car parks accommodated up to 1,500 vehicles. The terminal was connected to the new car parks via sheltered links. Once passengers arrive at the two-storey terminal building they pass through kiosks and automated options for self-check-in, self-bag tagging and self-bag drops. Their bags are transported to the aircraft via an advanced and automated baggage handling system. Similarly automated options, including face recognition technology, are used at immigration counters and departure gates. Biometric technology and 'fast and seamless travel' (FAST) services help to speed passenger throughput and increase efficiency. After security checks, passengers find themselves in 15,000 m^2 of shopping, dining and other retail spaces, featuring local, cultural and heritage-themed restaurants. This space also features a 300-metre-long central galleria, which is a glazed open space that visually connects the departure, check-in, arrival and transit areas across the terminal. The emphasis on the aesthetic appeal of the terminal is something that Changi considered important. It already boasts a butterfly garden, orchid and sunflower gardens as well as a koi pond.

The feelings of passengers using the terminal are an important part of its design. Mr Yam Kum Weng, Executive Vice President of CAG, one of the companies helping to develop the design for the new terminal, said: '*T4 breaks new ground in passenger experience for travellers, while ensuring smooth and efficient operations for airlines and airport agencies. Architecturally, the design of T4 will be functional, and yet have its own distinct character compared to the other three terminals at Changi Airport. Our focus for the development of T4 will be on its interior and ensuring that the design and layout continues to be passenger-centric and user-friendly. It will offer what passengers want – a good range of leisure amenities, convenient facilities and attractive commercial offerings.*' And with so many different companies involved in the day-to-day operation of the airport it was vital to include as many stakeholders as possible during the design. Workshops were conducted with various stakeholders, including airlines, ground handlers, immigration and security agencies, retail and food and beverage operators, as well as other users to ensure that the T4 design met the needs of each party. The objective was to ensure that T4, when operational, could deliver a seamless and refreshing experience for travellers, and also be a place where staff would feel proud and motivated to work.


John Seaton Callahan/Moment Unreleased/Getty Image

Why is process design important?

The case example 'Changi Airport' highlights a number of process design issues. The first is that the payback from good process design is clearly significant, or to reverse this logic, bad process design will quickly show in terms of cost and customer service. If you have been in an airport terminal when the baggage-handling system has malfunctioned, or security checks are taking longer than normal, or when it's just not coping with a holiday rush, you know the frustration of being the victim of poorly designed processes. And although competition between airports is perhaps not as fierce as between some other operations, in the long term poorly designed passenger, aircraft, cargo and baggage-handling processes will adversely affect any airport's viability. The case example also illustrates the importance of not being afraid to analyse processes at a very detailed level. This may include thoroughly understanding current processes so that any improvement can be based on the reality of what happens in practice. It will certainly involve allocating the tasks and associated capacity very carefully to appropriate parts of the process. And, for most processes, it will also involve a design that is capable of taking into consideration the variability that exists in most human tasks. These are the topics covered in this chapter.

OPERATIONS PRINCIPLE
Processes should always be designed to reflect customer and/or market requirements.

Processes also define how much of an operation's resource potential is realised. One cannot truly understand either the potential or the limitations of an operation's resources by simply listing them. Resources in themselves only add their full value if they are used effectively. It is the relationship between resources and the way they are used within processes that add value (hence the importance of the relationship between this and the previous chapter). Inadequate resources and disorganised processes make for poor operations, so it makes practical sense periodically to review their fitness. Does an operation have the right resources to provide an appropriate performance? What are, or will be, the main resource gaps now and in the future? Similar questions are useful when directed at processes. What are the key processes in the operation, what are processes good at, what are their weaknesses and/or vulnerabilities? How might processes cope with the ways in which the business might develop in the future?

Process networks

Processes should be designed in the context of the other processes that they are connected with. So it is worth revisiting the idea that we raised in Chapters 1 and 4 – the idea of 'networks'. In other words, how any operation is both made up of networks (of processes) and is a part of networks (of other operations). This idea is essential in making all networks, including process networks, work effectively. Figure 6.2 shows a simplified internal process network for one business. Through this network there are many 'process chains' – that is, a thread of processes within the network. And thinking about processes as part of a network has a number of advantages. First, understanding how and where a process fits into the internal network helps to establish appropriate objectives for the process. Second, one can check to make sure that everyone in a process has a clear 'line of sight' through to end customers, so that the people working in each process have a better chance of seeing how they contribute to satisfying the operation's customers. Even more important, one can ask the question, 'how can each process help the intermediate processes that lie between them and the customer to operate effectively?'. Third, a clear 'line of sight' backwards to the operation's suppliers makes the role and importance of suppliers easier to understand. Finally, reversing the question, not understanding how process chains interact can reduce the effectiveness of the whole operation, and increase the risk of disruption spreading. Process chains can become channels for disruption when things go wrong.

The very idea of 'process design' can be controversial

The whole idea of detailed process design can polarise opinion among operations managers. Two very opposing views are often expressed. But both can be simplistic. The first view is that everyone and everything is a process. And, if all work is a process, one should strive to apply a standardised set of disciplines and rules that can make what we do both more effective and more efficient. The objective, it is argued, is to find the best sequence of activities that describe

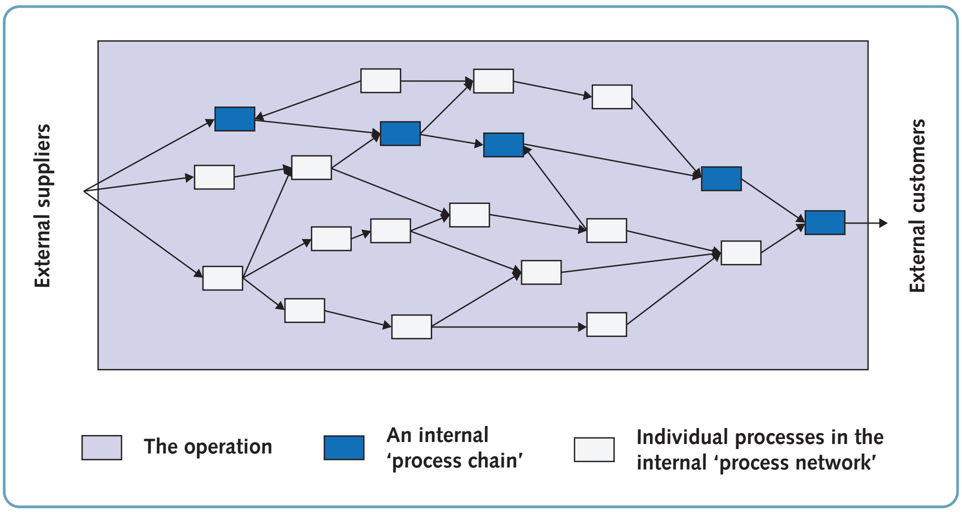


Figure 6.2 A process network within an operation showing an internal 'process chain'

the way things should be done, then preserve it in a formal set of rules that will ensure the same result 'every time, all the time'. The advantage of formal processes is the discipline it imposes. The opposing view believes that effective management must always rely on understanding, imagination and insight. You can't, according to this view, reduce everything to a standardised set of steps. People are the most important element in any processes, and people can't be treated like cogs in a machine. Processes, they say, describe everything as a machine-like set of mindless routine activities. But everything is about people, and effective management must always rely on the quality of people and how they are led with inspiration and imagination. Constraining all activities into a process straitjacket kills the essential humanity of working life.

> OPERATIONS PRINCIPLE
> *Not all processes are the same; some need to be highly specified, but not all. 'Process' simply means 'how we do things'.*

Of course, both of these rather stereotyped views are extreme, but they do represent the two ends of a spectrum of attitudes commonly expressed. A process is a set of resources and activities that produces value. That is exactly what we mean here by a process – no more and no less. Yes, everyone is part of a process. But that does not imply that all processes should be highly specified so jobs are deskilled, or that jobs should always be done in the same way. Some may need to be, but not all. The argument, sometimes referred to as 'people versus process', is based on a false distinction. People are the most important element in most processes, and 'process' is simply how we do things. In other words, not all processes are the same, nor should they be treated in the same way. 'Processes' need not consist of highly defined instructions. They can be looser than that. Described colloquially, processes are simply 'how you do things'. If appropriate, they could be defined in great detail, but they could also simply require adherence to a set of broad performance objectives and leave the details to someone's professional judgement (as in many professional service processes for example). The way processes are designed should depend on what we want them to do.

Worked example

Welcome to the lobby

'I know it's an emotive issue for many people, but lobbying is a perfectly reasonable, respectable and ethical activity. If you had responsibility for tens of thousands of employees, suppliers, and the millions who have their pension funds invested in companies like ours, wouldn't you want to put your point of view across? And we should be putting over our point of view as effectively as possible', said Jim Glazier, Head of Institutional and Public Affairs (IPA) for one of the largest pulp, paper and paper products companies in Europe. The European Parliament was a key target for interest groups. It maintained a register of accredited lobbyists who received special passes to access Members of the European Parliament (MEPs). But Jim was finding his team difficult. '*It's not that they are unskilled, they are good people who are vastly experienced. It's just that they are not used to being systematic or taking a standardised approach in doing their day-to-day jobs. We need to be smarter about how we go about our business.*' He argued that even skilled and experienced professionals like themselves needed to be more systematic about how they managed themselves. And the first step to systematising the team's activities would be to define all the processes that they engaged in.

Everyone looked, at best, sceptical. But it was Anders who was the first to object. '*I don't want to be obstructive, but I'm a qualified professional, not a process worker in a sausage factory. I don't just follow the rules, I use my brains and my experience to make vital decisions. I'm not simply a cog in a machine. And surely you don't want us to get into a box-ticking mentality? It just wouldn't work where we have constantly to think on our feet. Our job is the very antithesis of "process". Our activities are never routine or habitual, nor are they repetitive.*' Anders was warming to the theme now. '*Making them into "processes" would take away any creativity or intelligence; the job would become uninspiring and, well, just boring. It would impose unnecessary restrictions and constraints, reducing our freedom to act as we think appropriate. It would take away our humanity, even our self-worth!*'

Jim made an attempt to defuse the issue. '*Look*', he said, '*I know all of us, including me, would quite rightly resist any* *move that would reduce our ability to do our job to the best of our professional ability. I also know that when we hear the word "process" we immediately think about some of the plain stupid examples of "box-ticking" that we have all been subjected to. But that really isn't what I want to do – why would I? All I am proposing is that we should articulate three things: what we actually do, how we should go about doing it, and maybe how we could do it better. Is that such a radical proposal?*' Later he called in Marie Louise Bourg who headed the European Parliament lobbying operation. '*Do me a favour*', he asked, '*sit down with your team and simply write down what you do to prepare for, and carry out, your parliamentary briefing activities, and also, see how much difference there is between the ways each member of your team operates.*' Marie Louise was happy to do this. There was little to lose and, who knows, something useful could come of it!

At the next team meeting, Jim asked Marie Louise to present her findings. '*I will admit up front that the exercise was far more useful than I had anticipated; not so much for any great revelations, but for the serious questions it posed. First of all, we all had slightly different views of the various activities that we needed to complete, as well as their appropriate frequency, their correct sequence and their relative importance. It took over two hours just to agree on a relatively comprehensive list of tasks that needed to be completed quarterly, weekly and daily, never mind who should do them and in what order. We then discussed each stage in the process and tried to honestly assess how important it was in contributing to our mission, how good we were at doing it and how we could improve. It's only the beginning but I'm already sensing that the team are really looking forward to their ability to make some significant changes for the better.*'

Who should be involved in process design?

The decisions covered in the previous chapter would be taken at a relatively senior level. Deciding on operations layouts, the nature of the technology to be used, and the approach to employees' jobs, all have a clear strategic element. But who should be involved in the detailed analysis of processes?

In many (or most) organisations, process design is the province of some combination of technical experts and more junior staff. But there is a good case for regarding this as a mistake. As we implied earlier, process design is much more than a mere technical activity. It is important because it determines the actual performance of an operation's processes – and the performance of process networks makes up the performance of the total operation. Any operations manager surely must carry responsibility for how their processes are put together, which means that delegating total responsibility for the design of processes entirely to junior staff, or to internal 'experts' (although they have their role), is neglecting an important opportunity to improve performance. Certainly, operations managers will (and should) be held accountable if the processes go wrong. Of course, there is a difference between being engaged with, taking responsibility for and understanding process design, and getting bogged down in the minutiae.

6.2 Diagnostic question: *Are process performance objectives clear?*

The whole point of process design is to make sure that the performance of the process is appropriate for whatever it is trying to achieve. For example, if an operation competed primarily on its ability to respond quickly to customer requests, its processes would need to be designed to give fast throughput times. This would minimise the time between customers requesting a product or service and them receiving it. Similarly, if an operation competed on low price, then cost-related objectives are likely to dominate its process design. Some kind of logic should link what the operation as a whole is attempting to achieve and the performance objectives of its individual processes. This is illustrated in Table 6.1.

> **OPERATIONS PRINCIPLE**
> ***Process performance can be judged in terms of the levels of quality, speed, dependability, flexibility and cost they achieve.***

Table 6.1 The impact of strategic performance objectives on process design objectives and performance

Strategic performance objective	*Typical process design objectives*	*Some benefits of good process design*
Quality	• Provide appropriate resources, capable of achieving the specification of product or services • Error-free processing	• Products and service produced 'on-specification' • Less recycling and wasted effort within the process
Speed	• Minimum throughput time • Output rate appropriate for demand	• Short customer waiting time • Low in-process inventory
Dependability	• Provide dependable process resources • Reliable process output timing and volume	• On-time deliveries of products and services • Less disruption, confusion and rescheduling within the process
Flexibility	• Provide resources with an appropriate range of capabilities • Change easily between processing states (what, how, or how much is being processed)	• Ability to process a wide range of products and services • Low cost/fast product and service change • Low cost/fast volume and timing changes • Ability to cope with unexpected events (e.g. supply or processing failure)
Cost	• Appropriate capacity to meet demand • Eliminate process waste in terms of: – excess capacity – excess process capability – in-process delays – in-process errors – inappropriate process inputs	• Low processing costs • Low resource costs (capital costs) • Low delay/inventory costs (working capital costs)

In addition to the standard performance objectives, process designs are increasingly being judged on their environmental impact. The following example of Ecover illustrates one organisation with an ethical philosophy that stresses environmental impact objectives.

Case example

Ecover's ethical operation design[2]

Ecover is the world's largest producer of ecological detergents and eco-friendly cleaning and laundry products, with factories in France and Belgium that embody the company's commitment to sustainability. For example, Ecover factories operate entirely on green electricity – the type produced by wind generators, tidal generators and other natural sources. They also make the most of the energy they do use by choosing energy-efficient lighting. And, although the machinery they use in the factories is standard for the industry, they keep their energy and water consumption down by choosing low-speed appliances that don't require water to clean them. For example, the motors on their mixing machines can mix 25 tonnes of Ecover liquid while '*consuming no more electricity than* *a few flat irons*'. And they have a '*squeezy gadget that's so efficient at getting every last drop of product out of the pipes, they don't need to be rinsed through*'. Ecover say that they '. . . *hate waste, so we're big on recycling. We keep the amount of packaging used in our products to a minimum, and make sure that whatever cardboard or plastic we do use can be recycled, reused or refilled. It's an ongoing process of improvement; in fact, we've recently developed a new kind of green plastic we like to call "Plant-astic" that's 100 per cent renewable, reusable and recyclable – and made from sugarcane*.' Even the building is ecological, designed to follow the movement of the sun so that production takes place with the maximum amount of natural daylight (thus saving power). The factory's frame is built from pine, using bricks that are made from clay, wood pulp and mineral waste, requiring less energy to bake. The roofs are covered in a flowering plant that gives insulation all year round and is so effective that they don't need heating or air conditioning.

Kristoffer Tripplaar/Alamy Stock Photo

Process flow objectives

All the strategic performance objectives translate directly to process design, as shown in Table 6.1. But, because processes will be managed at a very operational level, process design also needs to consider a more 'micro' and detailed set of objectives. These are largely concerned with flow through the process. When whatever is being 'processed' (we shall refer to these as 'units' irrespective of what they are) enter a process they will progress through a series of activities where they are 'transformed' in some way. Between these activities the units may dwell for some time in inventories, waiting to be transformed by the next activity. This means that the time that a unit spends in the process (its throughput time) will be longer than the sum of all the transforming activities that it passes through. Also, the resources that perform the processes activities may not be used all the time, because not all units will necessarily require the same activities and the capacity of each resource may not match the demand placed upon it. So neither the units moving through the process, nor the resources performing the activities may be fully utilised. Because of this, the way that units leave the process is unlikely to be exactly the same as the way they arrive at the process. Figure 6.3 illustrates some of the 'micro' performance flow objectives that describe process flow performance and the process design factors that influence them. The flow objectives are:

- Throughput rate (or flow rate) – the rate at which units emerge from the process, i.e. the number of units passing through the process per unit of time.
- Throughput time – the average elapsed time taken for inputs to move through the process and become outputs.
- The number of units in the process (also called the 'work-in-process', or 'in-process inventory'), as an average over a period of time.
- The utilisation of process resources – the proportion of available time that the resources within the process are performing useful work.

OPERATIONS PRINCIPLE
Process flow objectives should include throughput rate, throughput time, work-in-process and resource utilisation, all of which are interrelated.

The design factors that will influence the flow objectives are the:

- variability of input arrival to the process;
- configuration of the resources and activities within the process;

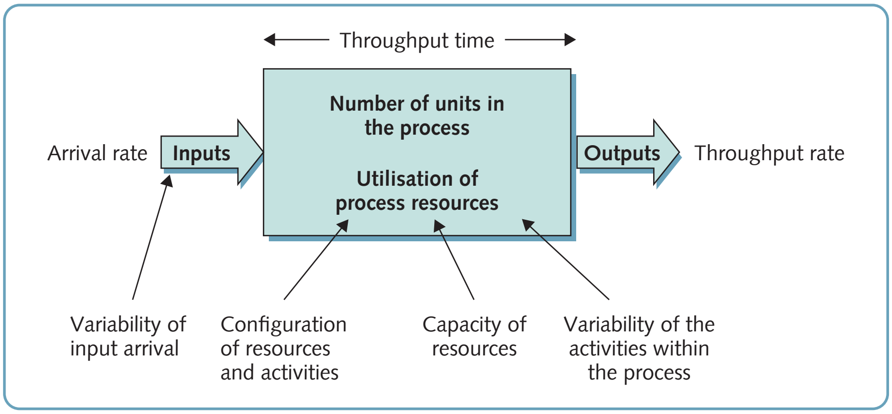


Figure 6.3 Micro process performance objectives and process design factors

- capacity of the resources at each point in the process; and
- variability of the activities within the process.

As we examine each of these design factors, we will be using a number of terms that, although commonly used within process design, need some explanation. These terms will be described in the course of the chapter, but Table 6.2 summarises them for reference.

Table 6.2 Some common process design terms

Process design term	*Description of term*
Process task	The sum of all the activities that must be performed by the process
Work content of the process task	The total amount of work within the process task measured in time units
Activity	A discrete amount of work within the overall process task
Work content of an activity	The amount of work within an activity measured in time units
Precedence relationship	The relationship between activities expressed in terms of their dependencies, i.e. whether individual activities must be performed before other activities can be started
Cycle time	The average time that the process takes between completions of units
Throughput rate	The number of units completed by the process per unit of time (= 1/cycle time)
Process stage	A work area within the process through which units flow; it may be responsible for performing several activities
Bottleneck	The capacity-constraining stage in a process; it governs the output from the whole process
Balancing	The act of allocating activities as equally as possible between stages in the process
Utilisation	The proportion of available time that the process, or part of the process, spends performing useful work
Starving	Underutilisation of a stage within a process caused by inadequate supply from the previous stage
Blocking	The inability of a stage in the process to work because the inventory prior to that stage is full
Throughput time	The elapsed time between a unit entering the process and its leaving the process
Queue time	The time a unit spends waiting to be processed

6.3 Diagnostic question: *How are processes currently designed?*

Existing processes are not always sufficiently well defined or described. Sometimes this is because they have developed over time without ever being formally recorded, or they may have been changed (perhaps improved) informally by the individuals who work in the process. But processes that are not formally defined can be interpreted in different ways, leading to confusion and inhibiting improvement. So, it is important to have some recorded visual descriptor of a process that can be agreed by all those who are involved in it. This is where process mapping comes in.

> OPERATIONS PRINCIPLE
> ***Process mapping is needed to expose the reality of process behaviour.***

Process mapping

Process mapping (or process blueprinting as it is sometimes called) at its most basic level involves describing processes in terms of how the activities within the process relate to each other. There are many, broadly similar, techniques that can be used for process mapping. However, all the techniques have two main features:

1. They identify the different types of activity that take place during the process.
2. They show the flow of materials or people or information through the process (or, put another way, the sequence of activities that materials, people or information are subjected to).

Different process mapping symbols are sometimes used to represent different types of activity. They can be arranged in order, and in series or in parallel, to describe any process. And although there is no universal set of symbols used all over the world, some are relatively common. Most derive either from the early days of 'scientific' management around a century ago, or, more recently, from information system flowcharting.

Worked example

Theatre lighting operation

Figure 6.4 shows one of the processes used in a theatre lighting operation. The company hires out lighting and stage effects equipment to theatrical companies and event organisers. Customers' calls are routed to the store technician. After discussing their requirements, the technician checks the equipment availability file to see if the equipment can be supplied from the company's own stock on the required dates. If the equipment cannot be supplied in-house, customers may be asked whether they want the company to try to obtain it from other possible suppliers. This offer depends on how busy and how helpful individual technicians are. Sometimes customers decline the offer and a 'Guide to Customers' leaflet is sent to the customer. If the customer does want a search, the technician will call potential suppliers in an attempt to find available equipment. If this is not successful the customer is informed, but if suitable equipment is located it is reserved for delivery to the company's site. If equipment can be supplied from the company's own stores, it is reserved on the equipment availability file and the day before it is required a 'kit wagon' is taken to the store where all the required equipment is assembled, taken back to the workshop and checked, and if any equipment is faulty it is repaired at this point. After that it is packed in special cases and delivered to the customer.

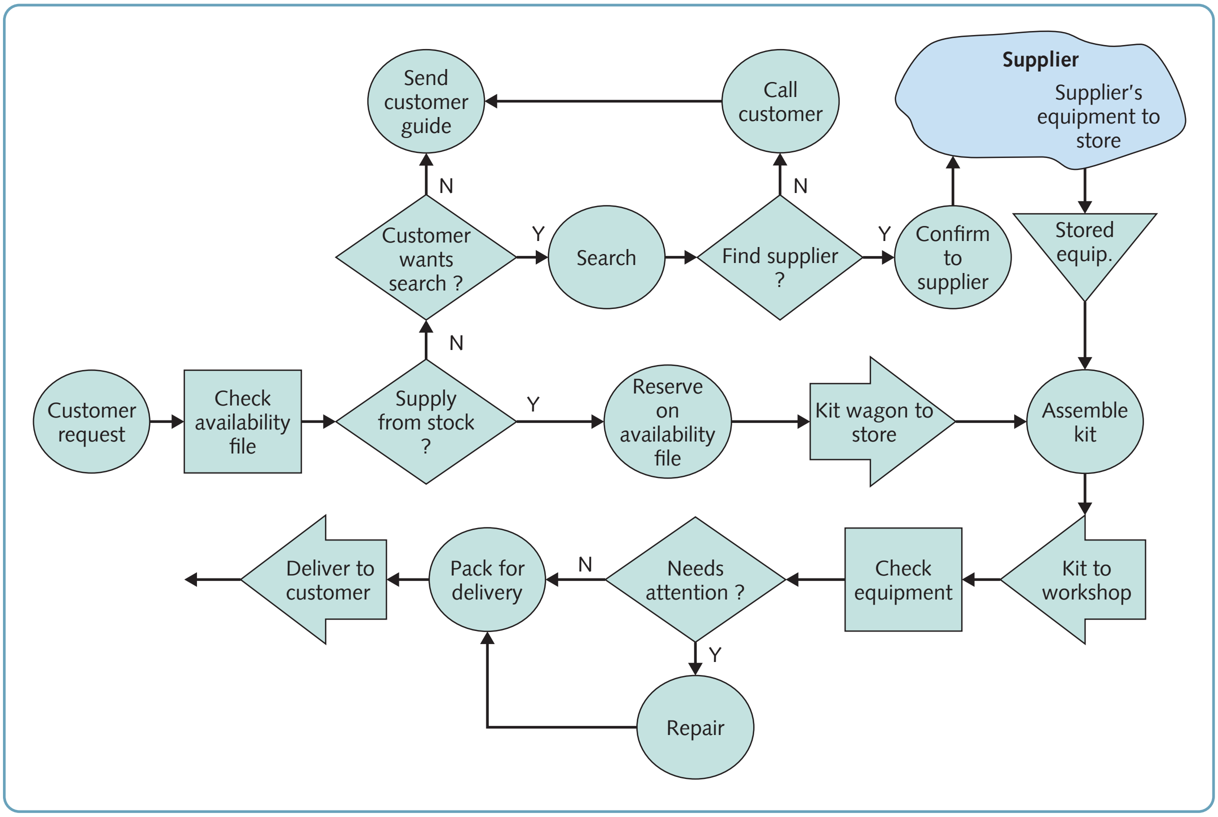


Figure 6.4 **Process map for 'enquiry-to-delivery' process at theatre lighting operation**

Different levels of process mapping

For a large process, drawing process maps at this level of detail can be complex. This is why processes are often mapped at a more aggregated level, called high-level process mapping, before more detailed maps are drawn. Figure 6.5 illustrates this for the total 'supply-and-install-lighting' process in the theatre lighting operation. At the highest level the process can be drawn simply as an input–transformation–output process with materials and customers as its input resources and lighting services as outputs. No details of how inputs are transformed into outputs are included. At a slightly lower or more detailed level, what is sometimes called an outline process map (or chart) identifies the sequence of activities but only in a general way. So the process of 'enquiry to delivery' that is shown in detail in Figure 6.4 is here reduced to a single activity. At the more detailed level in Figure 6.5, all the activities are shown in a 'detailed process map' (the activities within the process 'install and test' are shown).

Although not shown in Figure 6.5, an even more micro set of process activities could be mapped within each of the detailed process activities. Such a micro detailed process map could specify every single motion involved in each activity. Some quick-service restaurants, for example, do exactly that. In the lighting hire company example, most activities would not be mapped in any more detail than that shown in Figure 6.5. Some activities, such as 'return to base', are probably too straightforward to be worth mapping any further. Other activities, such as 'rectify faulty equipment', may rely on the technician's skills and discretion to the extent that the activity has too much variation and is too complex to map in detail. Some activities, however, may need mapping in more detail to ensure quality or to protect the company's interests. For

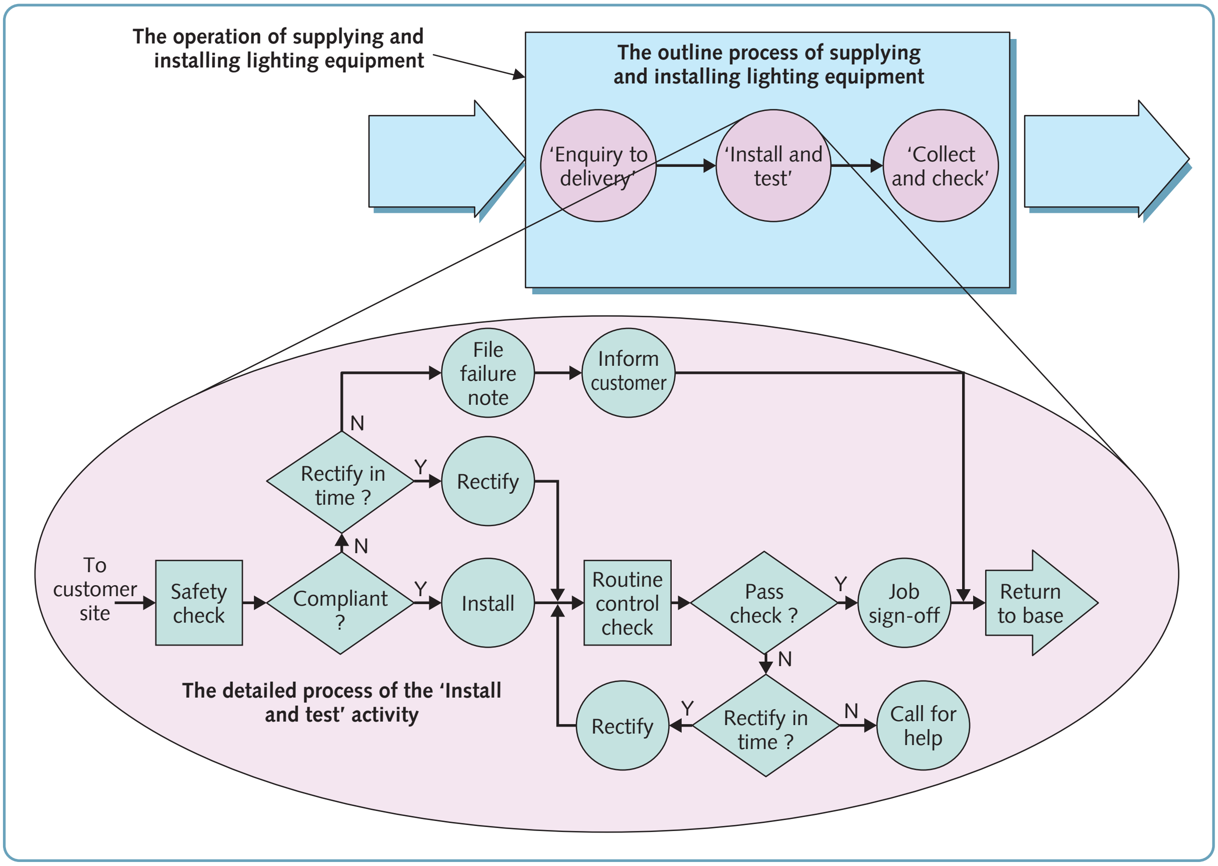


Figure 6.5 **The 'supply-and-install' operations process mapped at three levels**

example, the activity of safety-checking the customer's site to ensure that it is compliant with safety regulations will need specifying in some detail to ensure that the company can prove it exercised its legal responsibilities.

Process visibility

It is sometimes useful to map such processes in a way that makes the degree of visibility of each part of the process obvious. This allows those parts of the process with high visibility to be designed so that they enhance the customer's perception of the process. Figure 6.6 shows yet another part of the lighting equipment company's operation: the 'collect-and-check' process. The process is mapped to show the visibility of each activity to the customer. Here four levels of visibility are used. There is no hard and fast rule about this; many processes simply distinguish between those activities that the customer can see and those that they can't. The boundary between these two categories is often called the 'line of visibility'. In Figure 6.6 three categories of visibility are shown. At the very highest level of visibility, above the 'line of interaction', are those activities that involve direct interaction between the lighting company's staff and the customer. Other activities take place at the customer's site or in the presence of the customer but involve less or no direct interaction. Yet further activities (the two transport activities in this case) have some degree of visibility because they take place away from the company's base and are visible to potential customers, but are not visible to the immediate customer.

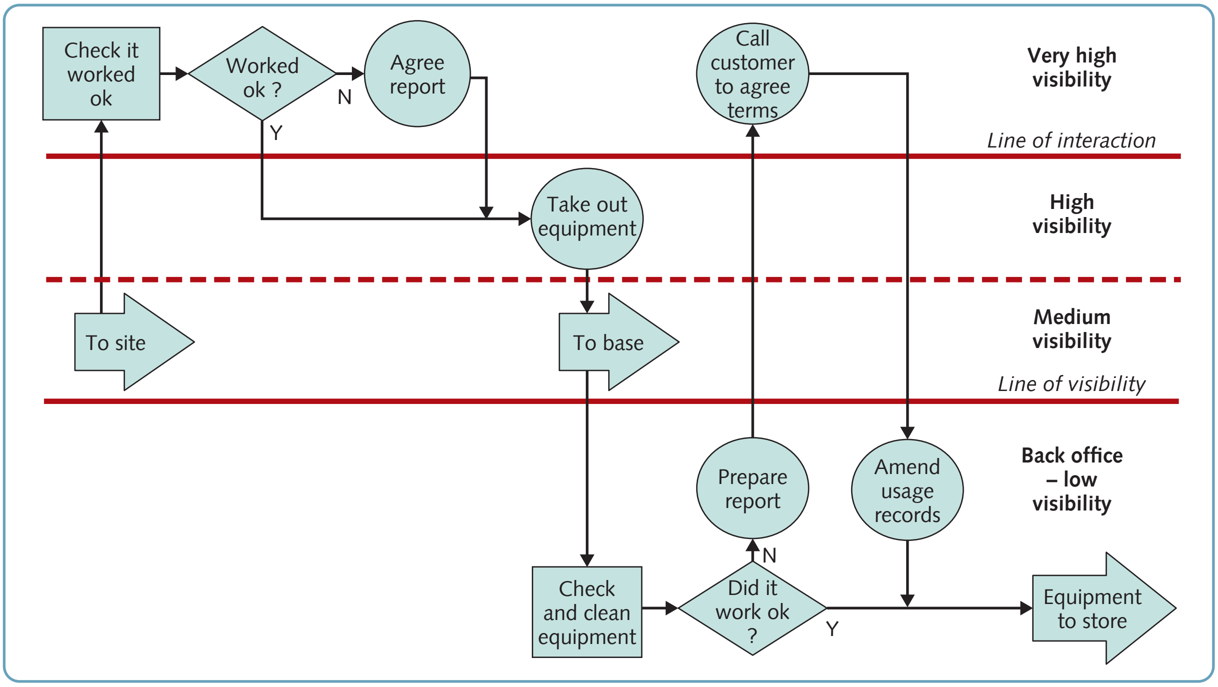


Figure 6.6 The 'collect-and-check' process mapped to show four different levels of process visibility

Visibility, customer experience and emotional mapping

Processes with a high level of customer 'visibility' cannot be designed in the same way as processes that deal with inanimate materials or information. 'Processing' people is different. As we discussed in Chapter 1, operations and processes that primarily 'transform' people present a particular set of issues. Material and information are processed, but customers experience the process. When a customer experiences a process it results in them feeling emotions, not all of which are necessarily rational. Most of us have been made happy, angry, frustrated, surprised, reassured or furious as customers in a process. Nor is the idea of considering how processes affect customer emotions confined to those processes that are intended to engage the emotions, as in entertainment-type operations such as theme parks, for example. Any high customer-contact product (or, more likely, service) always creates an experience for the customer. Moreover, customer experience will affect customer satisfaction, and therefore has the potential to produce customer loyalty, influence expectations and create emotional bonds with customers. This is why many service organisations perceive how customers experience their processes (the so-called 'customer journey') to be at the core of their process design.

Designing the customer experience

Designing processes with a significant experience content requires the systematic consideration of how customers may react to the experiences that the process exposes them to. This will include the sights, sounds, smells, atmosphere and general 'feeling' of the service. It is easy to confuse customer *service* and customer *experience*, but they are not interchangeable. Customer service refers to each single touchpoint with a service, while customer experience impacts feelings and emotion, and includes all service encounters throughout the entire customer journey. However, this does pose a problem. Because the experience is a personal interpretation of a service, it exists only in the customer's mind. As customers, our perceptions are individual to us and may not reflect any kind of 'objective' view of the actual situation; they are our own personal view of 'reality'. No two people can have the same experience. Some authorities suggest that experiences cannot be designed as such, only the mechanisms for creating the experience.[3] What operations managers can do, however, is to try to build-in the right clues and messages into their processes and try to design the customer's journey to create the right experience and emotions.

The concept of a 'servicescape', discussed in the previous chapter, is strongly related to consideration of engaging customers so that they connect with the process in a personal way. One of the most common methods of designing such processes is to consider what are commonly called 'touchpoints'. These have been described as 'everything the consumer uses to verify their service's effectiveness'.[4] They are the points of contact between a process and customers, and there might be many different touchpoints during the customer journey. It is the accumulation of all the experiences from every touchpoint interaction that shapes customers' judgement of the process. The features of a process at the touchpoints are sometimes called 'clues' (or 'cues'); these are the messages that customers receive or experience as they progress through the process. The emotions that result from these cues contain the messages that the customer will receive and therefore influence how a customer will judge the process.

Various tools and techniques have been developed to help design and improve the customer's journey and influence their experience. Most try to capture the series of touchpoints that customers encounter during a service – mapping the route customers take through websites, or clinical pathways in hospitals, for example. Doing this forces staff and managers to look at processes from the customer perspective. Another approach is to ask staff, managers or independent advisers to act as surrogate customers and conduct a 'walk-through audit'. Extending this idea, emotion mapping captures the emotions felt by customers as they move through the process. Both negative and positive emotions can be captured, discussed by managers and staff or the customer experience design team, changes made and then re-checked to see the impact of the changes.

When designing processes, managers need to ensure that all the messages coming from the clues at each stage of the process are consistent with the emotions they want the customers to experience and do not give them wrong or misleading messages about the process. In the same way as process mapping indicates the sequence and relationship between activities, so emotional mapping can indicate the types of emotions engendered in the customer's mind as they experience the process. Figure 6.7 shows how this might work for a visit to a clinic for a computed tomography (CT) scan. There are many ways that emotions can be mapped and different diagrammatic representations can be used. In this case, experiences are captured by asking what the patient is intended to, and actually, thinks, feels, says and does. From this, a simple scoring system has been used ranging from +3 (very positive) to –3 (very negative).

OPERATIONS PRINCIPLE
The design of processes that deal with customers should consider the emotions engendered at each stage of the process.

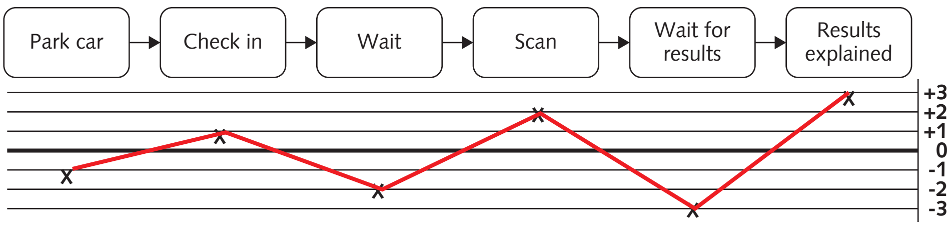
6.4 Diagnostic question: *Are process tasks and capacity configured appropriately?*

Process maps show how the activities of any particular process are currently arranged and help to suggest how they can be reconfigured. But there are also some general issues that must be understood before processes can be analysed. These relate to how the total task can be divided up within the process and determining how capacity is allocated. This, in turn, determines the flow through the process.

Getting to grips with process capacity means understanding the following issues:

- task precedence;
- series and parallel configurations;

What do you intend the patient to...						
Think	It's obvious where I park	This place looks friendly	Not too long to wait	This isn't too intimidating	Plenty to do while I wait	That is clearly explained
Feel	Not rushed	Comfortable, unconcerned	Relaxed	Unthreatened, relaxed	Not anxious	I understand what I am being told
Say	This is convenient	This seems a nice place	This seems a pleasant place	This is easier than I expected	It shouldn't be very long now	Thanks, these are my questions
Do	Park without hassle	Check in swiftly with no confusion	Sit down comfortably	Keep still during scan	Wait patiently	Respond to results appropriately



What does the patient actually ...						
Think	Where do I go?	I can see where to check in	How long will I have to wait?	This doesn't look unpleasant	How long will I have to wait?	I think I understand
Feel	Frustrated	Here at last	Uncertain	Unthreatened, relaxed	Frustrated, anxious	Reasssured
Say	Where is a parking space?	This is OK	What comes next?	This is easier than I expected	What is happening?	Thanks, these are my questions
Do	Drive round for a space	Confirm details	Look for something to read	Keep still during scan	Worry	Make appointment with doctor

Figure 6.7 Customer experience map of a visit for a computed tomography (CT) scan

- cycle time and process capacity – bottlenecks;
- process balancing;
- throughput, cycle time and work-in-process.

Task precedence

Any process redesign needs to preserve the inherent precedence of activities within the overall task. Task 'precedence' defines what activities must occur before others, because of the nature of the task itself. At its simplest level, task precedence is defined by:

- the individual activities that comprise the total process task; and
- the relationship between these individual activities.

Task precedence is usually described by using a 'precedence diagram', which, in addition to the above, also includes the following information:

> OPERATIONS PRINCIPLE
> *Process design must respect task precedence.*

- the time necessary to perform the total task (sometimes known as the 'work content' of the task); and
- the time necessary to perform each of the individual activities within the task.

Worked example

Computer repair service centre

A repair service centre receives faulty or damaged computers sent in by customers, repairs them and dispatches them back to the customer. Each computer is subject to the same set of tests and repair activities, and although the time taken to repair each computer will depend on the results of the tests, there is relatively little variation between individual computers.

Table 6.3 defines the process task of testing and repairing the computers in terms of the seven activities that comprise the total task, the relationship between the activities in terms of each activity's 'immediate predecessor', and the time necessary to perform each activity. Figure 6.8 shows the relationship between the activities graphically. This kind of illustration is called the 'precedence diagram' for the process task. It is useful because it indicates how activities must be sequenced in the eventual process design. For example, the process cannot perform activity 'b' before activity 'a' is completed. It does not determine how a process can be designed, but once the task has been analysed in this way, activities can be arranged to form the process's general configuration.

Table 6.3 Process task details for the 'computer test and repair' task

Activity code	*Activity name*	*Immediate predecessor*	*Activity time (mins)*
a	Preliminary test 1	–	5
b	Preliminary test 2	a	6
c	Dismantle	b	4
d	Test and repair 1	c	8
e	Test and repair 2	c	6
f	Test and repair 3	c	4
g	Clean/replace casing elements	d,e,f	10

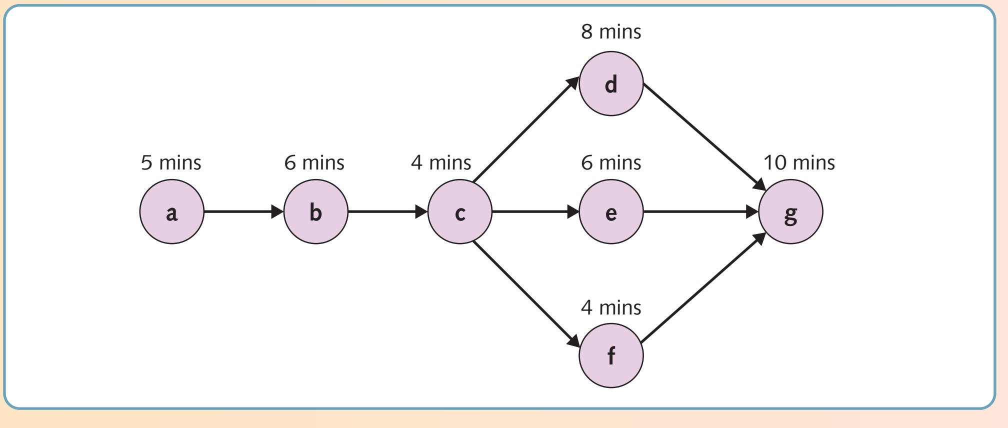


Figure 6.8 Precedence diagram showing the relationship between activities for the computer test and repair task

Series and parallel configurations

At its simplest level the general configuration of a process involves deciding the extent to which activities are arranged sequentially and the extent to which they are arranged in parallel.

For example, the task illustrated in Figure 6.8 involves seven activities that in total take 43 minutes. Demand is such that the process must be able to complete the test and repair task at the rate of one every 12 minutes in order to meet demand. One possible process design is to arrange the seven activities in a series arrangement of stages. The first question to address is, how many stages would this type of series arrangement require? This can be calculated by dividing the total work content of the task by the required cycle time.

In this case, number of stages = 43 minutes/12 minutes = 3.58 stages

Given the practical difficulties of having a fraction of a stage, this effectively means that the process needs four stages. The next issue is to allocate activities to each stage. Because the output from the whole process will be limited by the stage with most work (the sum of its allocated activities), each stage can have activities allocated to it up to a maximum allocation of 12 minutes. Figure 6.9 illustrates how this could be achieved. The longest stage (stage 2 in this case) will limit the output of the total process to one computer every 12 minutes and the other stages will be relatively under loaded.

However, there are other ways of allocating tasks to each stage, and involving the parallel arrangement of activities, that could achieve a similar output rate. For example, the four stages could be arranged as two parallel 'shorter' arrangements, with each stage performing approximately half of the activities in the total tasks. This is illustrated in Figure 6.10. It involves two two-stage arrangements, with stage 1 being allocated four activities that amount to 21 minutes of work and the second stage being allocated three activities that amount to 22 minutes of work. So, each arrangement will produce one repaired computer every 22 minutes (governed by the stage with the most work). This means that the two arrangements together will produce two repaired computers every 22 minutes, an average of one repaired computer every 11 minutes.

Loading each stage with more work and arranging the stages in parallel can be taken further. Figure 6.11 illustrates an arrangement where the whole test and repair task is performed at individual stages, all of which are arranged in parallel. Here, each stage will produce two repaired computers every 43 minutes and so together will produce four repaired computers every 43 minutes, an average output rate of one repaired computer every 10.75 minutes.

This simple example represents an important process design issue. Should activities in a process be arranged predominately in a single series 'long-thin' configuration, or predominantly in several 'short-fat' parallel configurations, or somewhere in-between? (Note that 'long' means the number of stages and 'fat' means the amount of work allocated to each stage.) Most processes will adopt a combination of series and parallel configurations, and in any particular situation there are usually technical constraints that limit either how 'long and thin' or how 'short and fat' the process can be. But there is usually a real choice to be made with a range of possible options. The advantages of each extreme of the long–thin to short–fat spectrum are very different and help to explain why different arrangements are adopted.

The advantages of the series-dominated (long–thin) configuration include:

- *A more controlled flow* through the process that is relatively easy to manage.
- *Simple materials handling* – especially if a product being manufactured is heavy, large or difficult to move.
- *Lower capital requirements*. If a specialist piece of equipment is needed for one element in the job, only one piece of equipment would need to be purchased; on short–fat arrangements every stage would need one.

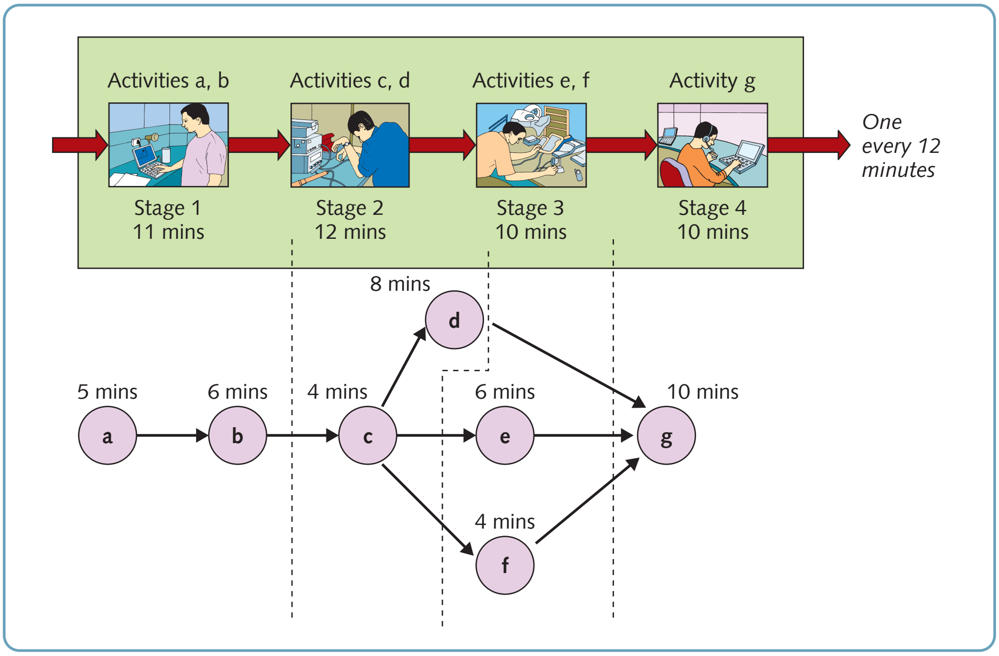


Figure 6.9 'Long-thin' arrangement of stages

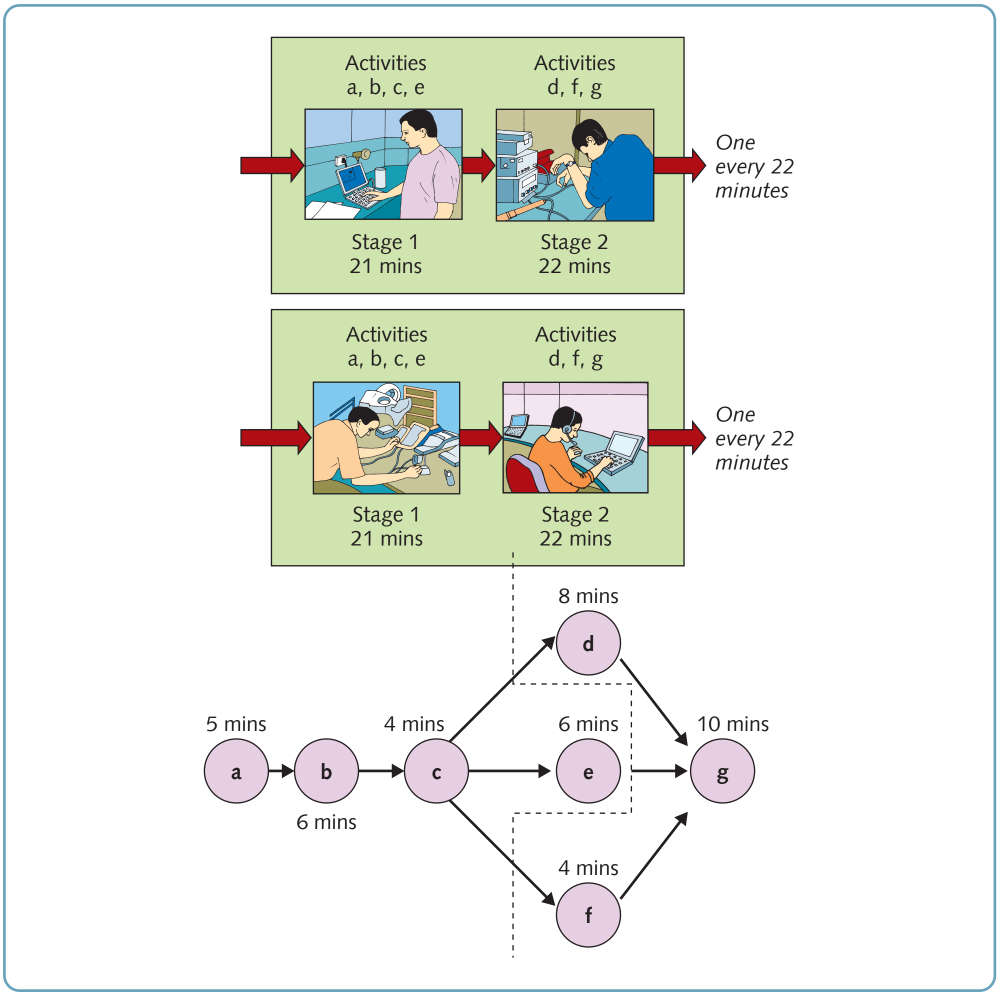


Figure 6.10 Intermediate configuration for the 'computer test and repair' task

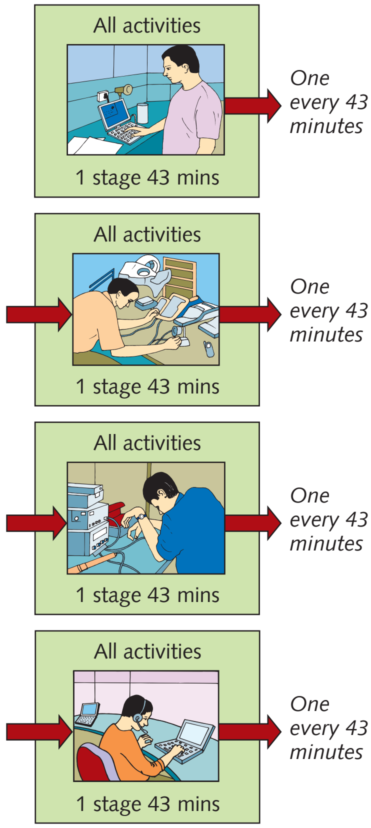


Figure 6.11 **'Short–fat' configuration of stages**

- *More efficient operation*. If each stage is performing only a small part of the total job, the person at the stage may have a higher proportion of direct productive work as opposed to the non-productive parts of the job, such as picking up tools and materials.

The advantages of the parallel-dominated (short–fat) configuration include:

- *Higher mix flexibility*. If the process needs to produce several types of product or service, each stage could specialise in different types.
- *Higher volume flexibility*. As volume varies, stages can simply be closed down or started up as required.
- *Higher robustness*. If one stage breaks down or ceases operation in some way, the other parallel stages are unaffected; a long–thin arrangement would cease operating completely.
- *Less monotonous work*. In the computer repair example, the staff in the short–fat arrangement repeat their tasks only every 43 minutes; in the long–thin arrangement it is every 12 minutes.

Cycle time and process capacity – bottlenecks

The cycle time of a process is the time between completed units emerging from it. Cycle time is a vital factor in process design and has a significant influence on most of the other detailed design decisions. It is usually one of the first things to be calculated because it can be used both

> **OPERATIONS PRINCIPLE**
> ***Process analysis derives from an understanding of the required process cycle time.***

to represent the demand placed on a process and the process's capacity. The cycle time also sets the pace or 'drum beat' of the process. However the process is designed, it must be able to meet its required cycle time. It is calculated by considering the likely demand for the products or services over a period and the amount of production time available in that period.

Process capacity

If the cycle time indicates the output that must be achieved by a process, the next decision must be how much capacity is needed by the process in order to meet the cycle time. To calculate this, a further piece of information is needed – the work content of the process task. The larger the work content of the process task and the smaller the required cycle time, the more capacity will be necessary if the process is to meet the demand placed upon it.

Process balancing

Balancing a process involves attempting to allocate activities to each stage as equally as possible. Because the cycle time of the whole process is limited by the longest allocation of activity times to an individual stage, the more equally work is allocated the less time will be 'wasted' at the other stages in the process. In practice it is nearly always impossible to achieve perfect balance, so some degree of imbalance in work allocation between stages will occur. The effectiveness of the balancing activity is measured by balancing loss. This is the time wasted through the unequal allocation of activities as a percentage of the total time invested in processing. The stage of a process that has the longest allocation of work is known as the bottleneck. It is the bottleneck stage that will dictate the output of the whole process. This is illustrated in Figure 6.12. Here the computer test and repair task is used to illustrate balancing loss for the

Worked example

Passport office

Suppose the regional government office that deals with passport applications is designing a process that will check applications and issue the documents. The number of applications to be processed is 1,600 per week and the time available to process the applications is 40 hours per week.

$$\text{Cycle time for the process} = \frac{\text{time available}}{\text{number to be processed}}$$

$$= \frac{40}{1600} = 0.025 \text{ hours}$$

$$= 1.5 \text{ minutes}$$

So the process must be capable of dealing with a completed application once every 1.5 minutes, or 40 per hour.

For the passport office, the total work content of all the activities that make up the total task of checking, processing and issuing a passport is, on average, 30 minutes.

So, a process with one person would produce a passport every 30 minutes. That is, one person would achieve a cycle time of 30 minutes. Two people would achieve a cycle time of 30/2 = 15 minutes, and so on.

Therefore, the general relationship between the number of people in the process (its capacity in this simple case) and the cycle time of the process is:

$$\frac{\text{Work content}}{N} = \text{cycle time}$$

Where: N = *the number of people in the process*

Therefore, in this case:

$$N = \frac{30}{\text{Cycle time}}$$

In this case:

$$N = \frac{30}{1.5} = 20 \text{ people}$$

So, the capacity that this process needs if it is to meet demand is 20 people.

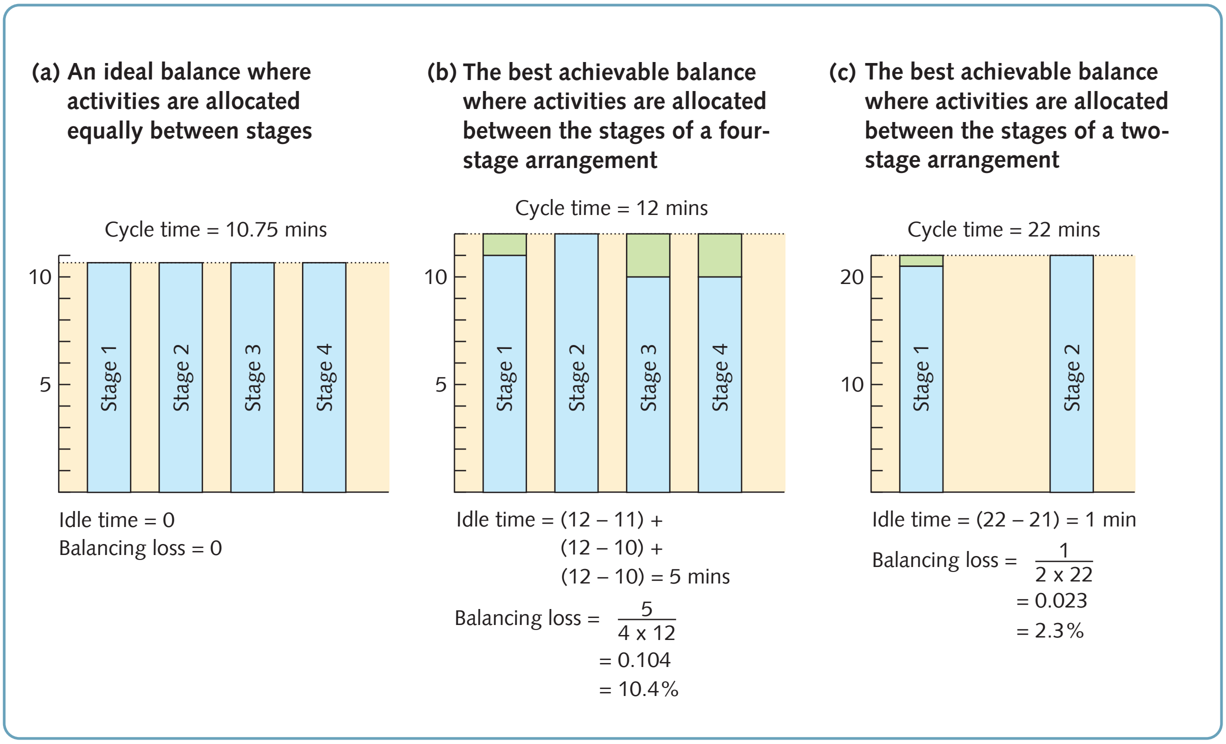


Figure 6.12 **Balancing loss is that proportion of the time invested in processing the product or service that is not used productively**

> OPERATIONS PRINCIPLE
> ***Allocating work equally to each stage in a process (balancing) smooths flow and avoids bottlenecks.***

'long–thin' arrangement of four sequential stages and the 'intermediate' arrangement of two parallel two-stage arrangements.

Figure 6.12(a) shows the ideal allocation of activities with each stage perfectly balanced. Here exactly a quarter of the total work content (10.75 minutes) has been allocated to each of the four stages. Every 10.75 minutes each stage performs its activities and passes a computer on to the next stage, or out of the process in the case of stage 4. No stage suffers any idle time and, because the stages are perfectly balanced, balancing loss = 0. However, because of the actual times of each activity, it is not possible to equally allocate work to each stage. Figure 6.12(b) shows the best allocation of activities. Most work is allocated to stage 2, so that stage will dictate the cycle time of the whole process. Stage 1 has only 11 minutes of work and so will be idle for (12−11) = 1 minute every cycle (or alternatively will keep processing one computer every 11 minutes and the build-up of inventory between stage 1 and stage 2 would grow to infinity). Similarly, stages 3 and 4 have idle time, in this case both have (12−10) = 2 minutes idle time. They can only process one computer every 12 minutes because stage 2 will only pass forward a computer to them every 12 minutes. So, they are being starved of work for 2 minutes every 12 minutes. In practice, stages that are not the bottleneck stage may not actually be idle for a period of time every cycle. Rather they will slow down the pace of work to match the time of the bottleneck stage. Nevertheless, this is still effectively idle time because under conditions of perfect balance they could be performing useful work.

So, every cycle, all four stages are investing an amount of time equivalent to the cycle time to produce one completed computer. The total amount of invested time therefore is the number of stages in the process multiplied by the cycle time. In this case, total invested time = 4 × 12 = 48 minutes. The total idle time for every computer processed is the sum of the idle times at the non-bottleneck stages, in this case 5 minutes.

Balancing loss is the amount of idle time as a percentage of the total invested time. In this case, balancing loss = 5/(4 × 12) = 0.104 = 10.4%.

Figure 6.12(c) makes the same calculation for the intermediate process described earlier. Here too, two-stage arrangements are placed in parallel. Stage 2 has the greatest allocation of work at 22 minutes, and will therefore be the bottleneck of the process. Stage 1 has 21 minutes' worth of work and therefore one minute of idle time every cycle. The total invested time in the process each cycle = 2 × 22. So, the balancing loss = 1/(2 × 22) = 0.023 = 2.3%.

Case example

London Underground tackles a bottleneck[5]

Anyone who has travelled on a busy mass transport system like the London Underground knows how crowded it can be, often with queues of passengers building up at various points as they move to or from their trains. One point that can become a bottleneck for passengers on the London Underground is the escalators. Traditionally, in London, passengers stand on the right side of an escalator, leaving the left side free for those who want to walk up or down. But in an attempt to reduce the bottleneck at the escalators, Transport for London, who run the system, trialled a new arrangement that they believed would increase the capacity of its escalators at the Holborn station. Building new stations is expensive, so any way of increasing the capacity of existing ones is going to be attractive, and Holborn is a particularly busy station. The new (and radical, for Londoners) arrangement was to instruct passengers at peak times not to walk, but to stand on both sides of the escalator. The decision was also based on the fact that the escalators at Holborn are over 24 metres high. Apparently, height makes a big difference to the willingness of passengers to walk up escalators. When they are only a few metres high, most people will walk up them. At 30 metres, only the very energetic will. As shown in Figure 6.13, the trial was technically successful in that capacity increased significantly. However, the experiment was not made permanent. Why? Apparently it offended two aspects of human behaviour. First, it slowed the (vocal) minority of people who wanted to race up the escalator as their gym workout for the day. Second, it took away the feeling that travellers had at least some degree of choice (even if most chose not to exercise it).

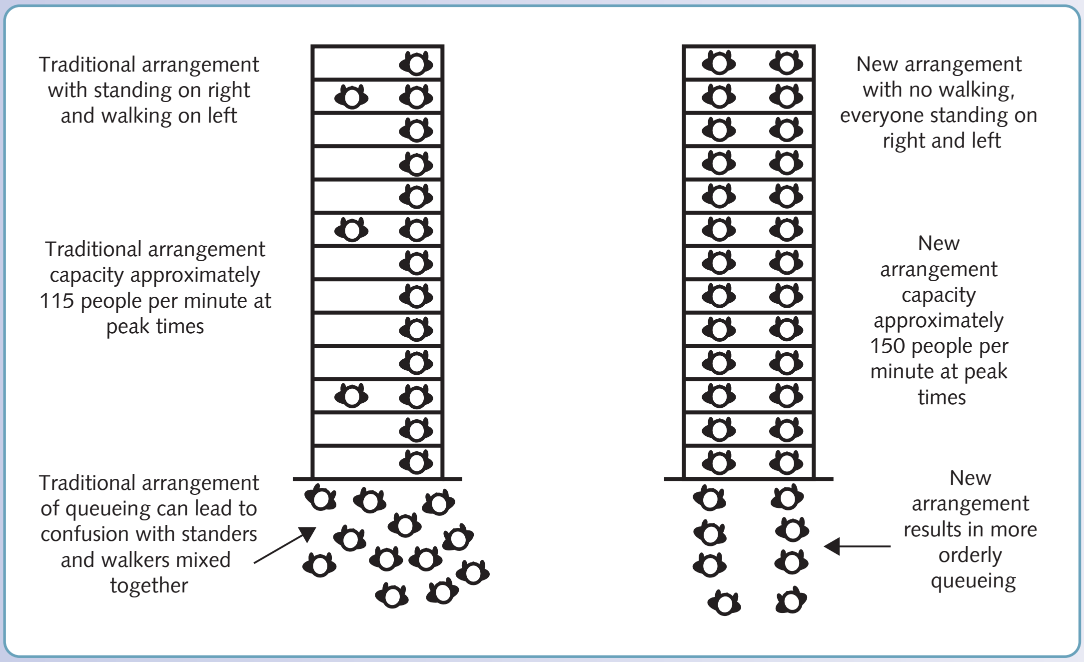


Figure 6.13 Requiring passengers to stand on both sides of the escalator makes it less of a bottleneck

Throughput, cycle time and work-in-process

The cycle time of a process is a function of its capacity. For a given amount of work content in the process task, the greater the capacity of the process the smaller its cycle time. In fact, the capacity of a process is often measured in terms of its cycle time, or more commonly the reciprocal of cycle time that is called 'throughput rate'. So, for example, a theme park ride would be described as having the capacity of 1,000 customers an hour, or an automated bottling line as having a capacity of 100 bottles a minute, and so on. However, a high level of capacity (short cycle time and fast throughput rate) does not necessarily mean that material, information or customers can move quickly through the process. This will depend on how many other units are contained within the process. If there is a large number of units within the process they may have to wait in 'work-in-process' inventories for part of the time they are within the process (throughput time).

Little's Law

The mathematical relationship that relates cycle time to work-in-process and throughput time is called Little's Law. It is simple but very useful, and it works for any stable process.

Little's Law can be stated as:

$$\text{Throughput time} = \text{Work-in-process} \times \text{Cycle time}$$

Or:

$$\text{Work-in-process} = \text{Throughput time} \times (1/\text{Cycle time})$$

That is:

$$\text{Work-in-process} = \text{Throughput time} \times \text{Throughput rate}$$

For example, in the case of the computer test and repair process with four stages:

$$\text{Cycle time} = 12 \text{ minutes (loading on the bottleneck station)}$$

$$\text{Work-in-process} = 4 \text{ units (one at each stage of the process assuming there is no space for inventory to build up between stages)}$$

$$\text{Therefore, Throughput time} = \text{Work-in-process} \times \text{Cycle time} = 12 \times 4 = 48 \text{ minutes}$$

> **OPERATIONS PRINCIPLE**
> ***Little's Law states that throughput time = work-in-process × cycle time.***

Similarly, for the worked example of the passport office earlier, suppose the office has a 'clear desk' policy that means that all desks must be clear of work by the end of the day. How many applications should be loaded onto the process in the morning in order to ensure that every one is completed and desks are clear by the end of the day?

From before: Cycle time = 1.5 minutes, and assuming 7.5 hour (450-minute) working day, from Little's Law:

$$\text{Throughput time} = \text{Work-in-process} \times \text{Cycle time}$$

$$450 \text{ minutes} = \text{Work-in-process} \times 1.5$$

$$\text{Therefore} \quad \text{Work-in-process} = 450/1.5 = 300$$

So, 300 applications can be loaded onto the process in the morning and be cleared by the end of the working day.

Worked example

Little's Law at a seminar

Mike was totally confident in his judgement: '*You'll never get them back in time*', he said. '*They aren't just wasting time, the process won't allow them to all have their coffee and get back for 11 o'clock.*' Looking outside the lecture theatre, Mike and his colleague Shaun were watching the 20 businessmen who were attending the seminar queuing to be served coffee and biscuits. The time was 10.45 and Shaun knew that unless they were all back in the lecture theatre at 11 o'clock there was no hope of finishing his presentation before lunch. '*I'm not sure why you're so pessimistic*', said Shaun. '*They seem to be interested in what I have to say and I think they will want to get back to hear how operations management will change their lives.*' Mike shook his head. '*I'm not questioning their motivation*', he said, '*I'm questioning the ability of the process out there to get through them all in time. I have been timing how long it takes to serve the coffee and biscuits. Each coffee is being made fresh and the time between the server asking each customer what they want and them walking away with their coffee and biscuits is taking 48 seconds. Remember that, according to Little's Law, throughput equals work-in-process multiplied by cycle time. If the work-in-process is the 20 managers in the queue and cycle time is 48 seconds, the total throughput time is going to be 20 multiplied by 0.8 minutes which equals 16 minutes. Add to that sufficient time for the last person to drink their coffee and you must expect a total throughput time of a bit over 20 minutes. You just haven't allowed long enough for the process.*' Shaun was impressed. '*Er. . . what did you say that law was called again?*' '*Little's Law*', said Mike.

Worked example

Little's Law at an IT support unit

Every year it was the same. All the workstations in the building had to be renovated (tested, new software installed, etc.) and there was only one week in which to do it. The one week fell in the middle of the August vacation period when the renovation process would cause minimum disruption to normal working. Last year the company's 500 workstations had all been renovated within one working week (40 hours). Each renovation last year took on average 2 hours, and 25 technicians had completed the process within the week. This year there would be 530 workstations to renovate but the company's IT support unit had devised a faster testing and renovation routine that would only take on average 1.5 hours instead of 2 hours. How many technicians would be needed this year to complete the renovation processes within the allotted week?

Last year:

Work-in-process (WIP) = 500 workstations

Time available (T_t) = 40 hours

Average time to renovate = 2 hours

Therefore, Throughput rate (T_r) = 1/2 per technician

= $0.5N$

Where N = number of technicians

Little's Law:

$$\text{WIP} = T_t \times T_r$$

$$500 = 40 \times 0.5\,N$$

$$= 25 \text{ technicians}$$

This year:

Work-in-process (WIP) = 530 workstations

Time available = 40 hours

Average time to renovate = 1.5 hours

Throughput rate (T_r) = 1/1.5 per technician

= $0.67N$

Where N = number of technicians

Little's Law:

$$\text{WIP} = T_t \times T_r$$

$$530 = 40 \times 0.67N$$

$$= 19.88 \text{ technicians}$$

6.5 Diagnostic question: *Is process variability recognised?*

So far in our treatment of process analysis we have assumed that there is no significant variability either in the demand to which the process is expected to respond, or in the time taken for the process to perform its various activities. Clearly, this is not the case in reality. So, it is important to look at the variability that can affect processes and take account of it. However, do not dismiss the deterministic analysis we have been examining up to this point. At worst it provides a good first approximation to analysing processes, while at best, the relationships that we have discussed do hold for average performance values.

Sources of variability in processes

There are many reasons why variability occurs in processes. A few of these possible sources of variation include:

- The late (or early) arrival of material, information or customers at a stage within the process.
- The temporary malfunction or breakdown of process technology within a stage of the process.
- The necessity for recycling 'misprocessed' materials, information or customers to an earlier stage in the process.
- The misrouting of material, information or customers within the process that then needs to be redirected.
- Each product or service being processed might be different – for example, different types of enquiries at a utility's contact centre.
- Products or services, although essentially the same, might require slightly different treatment. For instance, in the computer test and repair process, the time of some activities will vary depending on the results of the diagnostic checks.
- With any human activity there are slight variations in the physical coordination and effort on the part of the person performing the task that result in variation in activity times, even of routine activities.

All these sources of variation within a process will interact with each other, but result in two fundamental types of variability:

- Variability in the demand for processing at an individual stage within the process, usually expressed in terms of variation in the inter-arrival times of units to be processed.
- Variation in the time taken to perform the activities (i.e. process a unit) at each stage.

Activity time variability

The effects of variability within a process will depend on whether the movements of units between stages, and hence the inter-arrival times of units at stages, are synchronised or not. For example, consider the computer test and repair process described previously. Figure 6.14 shows the average activity time at each stage of the process, but also the variability around the average time. Suppose that it was decided to synchronise the flow between the four stages by using an indexing conveyor or a simple traffic light system that ensured all movement between the stages happened simultaneously. The interval between each synchronised movement would have to be set at an interval that would allow all stages to have finished their activities

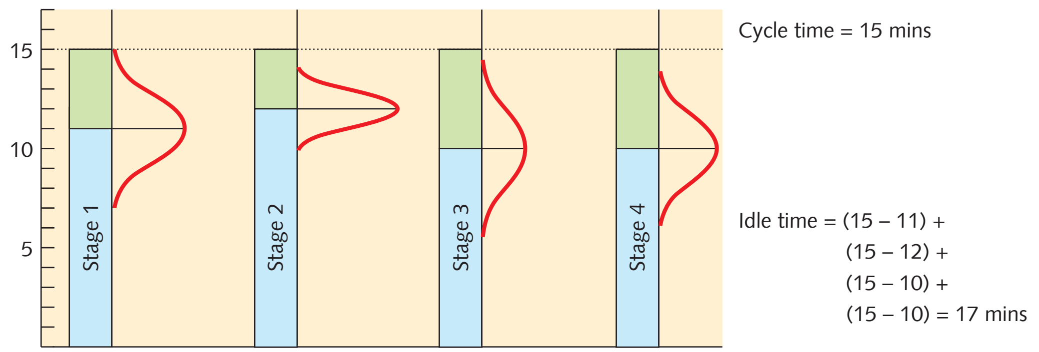


Figure 6.14 Processing time variability in a synchronised process; cycle time will need to accommodate the longest activity time at any of the stages

irrespective of whether they had experienced a particularly fast or particularly slow activity time. In this case from Figure 6.14, that synchronised indexing time would have to be set at 15 minutes. This then becomes the effective cycle time of the process. Note that the effective bottleneck stage is now stage 1 rather than stage 2. Although stage 2 has the longer average activity time (12 minutes), stage 1 with an average activity time of 11 minutes has a degree of variability that results in a maximum activity of 15 minutes. Note also that every stage will experience some degree of idle time, the average idle time at each station being the cycle time minus the average activity time at that station. This reduction in the efficiency of the process is only partly a result of its imbalance. The extra lost time is as a result of activity time variability.

> **OPERATIONS PRINCIPLE**
> *Variability in a process acts to reduce its efficiency.*

This type of effect is not at all uncommon. For example, automobiles are assembled using a moving belt assembly line whose speed is set to achieve a cycle time that can accommodate activity time variability. However, a more common arrangement, especially when processing information or customers, is to move units between stages as soon as the activities performed by each stage are complete. Here, units move through the process in an unsynchronised manner rather than having to wait for an imposed movement time. This means that each stage may spend less time waiting to move its unit forward, but it does introduce more variation in the demand placed on subsequent stations. When movement is synchronised, the inter-arrival time of units at each stage is fixed at the cycle time. Without synchronisation, the inter-arrival time at each stage will itself be variable.

Arrival time variability

To understand the effect of arrival variability on process performance it is first useful to examine what happens to process performance in a very simple process as arrival time changes under conditions of no variability. For example, the simple process shown in Figure 6.15 comprises one stage that performs exactly 10 minutes of work. Units arrive at the process at a constant and predictable rate. If the arrival rate is one unit every 30 minutes, then the process will be utilised for only 33.33 per cent of the time, and the units will never have to wait to be processed. This is shown as point A in Figure 6.15. If the arrival rate increases to one arrival every 20 minutes, the utilisation increases to 50 per cent, and again the units will not have to wait to be processed. This is point B in Figure 6.15. If the arrival rate increases to one arrival every 10 minutes, the process is now fully utilised, but, because a unit arrives just as the previous one has finished being processed, no unit has to wait. This is point C in Figure 6.15. However, if the arrival rate ever exceeded one unit every 10 minutes, the waiting line in front of the process

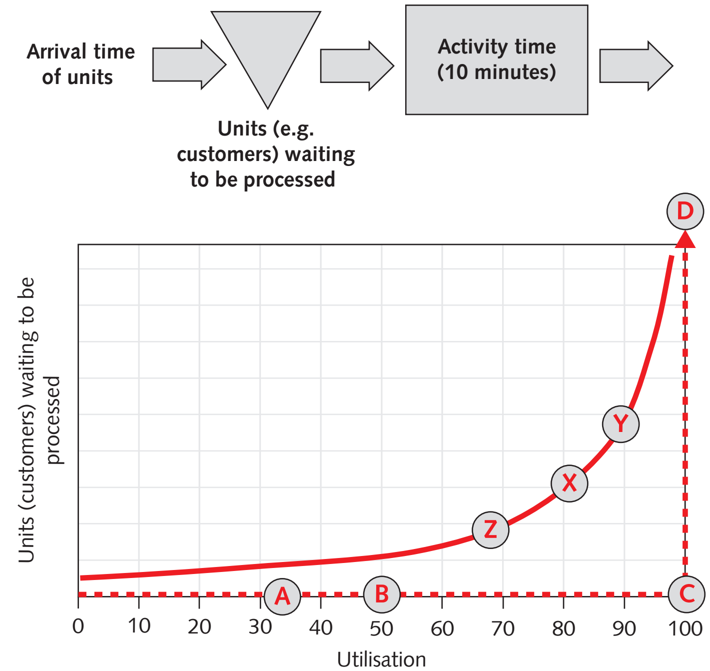


Figure 6.15 **The relationship between process utilisation and number of units waiting to be processed**

> OPERATIONS PRINCIPLE
> ***Process variability results in simultaneous waiting and resource underutilisation.***

activity would build up indefinitely, as is shown as point D in Figure 6.15. So, in a perfectly constant and predictable world, the relationship between process waiting time and utilisation is a rectangular function–as shown by the dotted line in Figure 6.15.

Activity and arrival time variability together

The dotted line in Figure 6.15 is, of course, purely theoretical. It is difficult to imagine any real process with absolutely zero variation. In practice there is likely to be a combination of activity and arrival time variation. Imagine a checkout station in a supermarket. Sometimes it will be idle with no customers waiting, at other times it will have several customers waiting. If one recorded the activity and number of waiting customers, a typical result could be point X in Figure 6.15, with both some underutilisation and a net positive average queue. Now suppose that the supermarket manager was unhappy because point X represented too low a level of utilisation; she might close some checkout stations. Average utilisation would increase, but average customer waiting time is also likely to increase, say to point Y in Figure 6.15. Conversely, if the manager is unhappy with point X because customers are waiting too long, she might react by opening more checkout stations. Average customer waiting time would probably decrease, but so would utilisation to, say, point Z in Figure 6.15.

Generally, as a process moves closer to 100 per cent utilisation the higher the average waiting time will become. Or, to put it another way, because of variability, the only way to guarantee very low waiting times for the units is to suffer low process utilisation. In fact, the greater the variability, the more the waiting time utilisation deviates from the simple rectangular function of the 'no variability' conditions (the dotted line in Figure 6.15). A set of curves for a typical process is shown in Figure 6.16(a). This phenomenon, which is sometimes called 'the OM triangle', has important implications for the design of processes. In effect it presents three

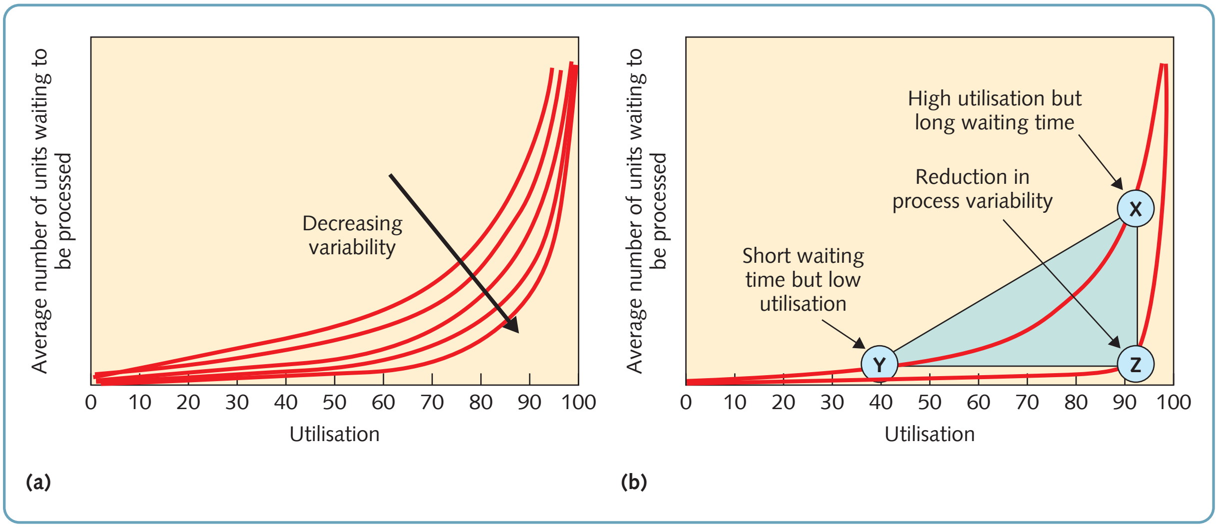


Figure 6.16 The relationship between process utilisation and number of units waiting to be processed for variable arrival and activity times

options to process designers wishing to improve the waiting time or utilisation performance of their processes, as shown in Figure 6.16(b). Either:

- accept long average waiting times and achieve high utilisation (point X); or
- accept low utilisation and achieve short average waiting times (point Y); or
- reduce the variability in arrival times, activity times, or both, and achieve higher utilisation and short waiting times (point Z).

To analyse processes with both inter-arrival and activity time variability, queuing or 'waiting line' analysis can be used. This is treated in the supplement to this chapter. But, do not dismiss the relationship shown in Figures 6.15 and 6.16 as some minor technical phenomenon. It is far more than this. It identifies an important choice in process design that could have strategic implications. Which is more important to a business, fast throughput time, or high utilisation of its resources? The only way to have both of these simultaneously is to reduce variability in its processes, which may itself require strategic decisions such as limiting the degree of customisation of products or services, or imposing stricter limits on how products or services can be delivered to customers, and so on. It also demonstrates an important point concerned with the day-to-day management of process – the only way to absolutely guarantee 100 per cent utilisation of resources is to accept an infinite amount of work-in-process and/or waiting time. We will take this point further in Chapter 8 when we deal with capacity management.

OPERATIONS PRINCIPLE
Process design involves some choice between utilisation, waiting time and variability reduction.

Critical commentary

- Unless you are one of those who instinctively reacts against the very idea of 'processes', there is not too much that would be considered contentious in this chapter. However, some practitioners would reject the idea of mapping processes as they exist currently. Rather, they would advocate a more radical 'clean sheet of paper' approach. Only by doing this, they would say, could one be sufficiently imaginative in the redesign of processes. Having said that the details of process analysis are not too contentious, there is one criticism that challenges the whole basis of how we think about processes. All the processes used as illustrations in this chapter involve relatively simple stage-by-stage arrangements very much like an 'assembly line' where work is linearly synchronised to deliver a completed task. Yet many business processes are far more complex than this with fuzzy and simultaneous activities that interact in unpredictable ways. And, although we have made the point that not all processes should be designed in the same way, there is relatively little guidance in conventional process design theory about how to tackle these more complex processes.

SUMMARY CHECKLIST

- ☐ Have a clear set of performance objectives for each process been set?
- ☐ Do the process design objectives clearly relate to the business's strategic objectives?
- ☐ Is the following information known for all key processes in the operation?
  - The throughput or flow rate of the process?
  - The throughput time of the process?
  - The number of units in the process (work-in-process)?
  - The utilisation of process resources?
- ☐ Are processes documented using process mapping techniques?
- ☐ Are the emotional responses of any customers/people who are being 'processed' understood?
- ☐ Are formal process descriptions followed in practice?
- ☐ If not, should the process descriptions be changed or should existing process descriptions be enforced?
- ☐ Is it necessary for process descriptions to include the degree of visibility at each stage of the process?
- ☐ Are the details of task precedence known for each process?
- ☐ Have the advantages and disadvantages of series and parallel configurations been explored?
- ☐ Is the process balanced? If not, can the bottleneck stages be redesigned to achieve better balance?
- ☐ Are the relationships between throughput, cycle time and work-in-process understood (Little's Law)?
- ☐ Are the sources of process variability recognised?
- ☐ Has the effect of variability been recognised in the design of the process?

Case study

The Action Response Applications Processing Unit (ARAPU)

Introduction

Action Response is a London-based charity dedicated to providing fast responses to critical situations throughout the world. It was founded by Susan N'tini, its Chief Executive, to provide relatively short-term aid for small projects until they could obtain funding from larger donors. The charity receives requests for cash aid usually from an intermediary charity and looks to process the request quickly, providing funds where and when they are needed. '*Give a man a fish and you feed him today, teach him to fish and you feed him for life, it's an old saying and it makes sense but, and this is where Action Response comes in, he might starve while he's training to catch fish*.' (Susan N'tini)

Nevertheless, Susan does have some worries. She faces two issues in particular. First, she is receiving complaints that funds are not getting through quickly enough. Second, the costs of running the operation are starting to spiral. She explains: '*We are becoming a victim of our own success. We have striven to provide greater accessibility to our funds; people can access application forms via the internet, by post and by phone. But we are in danger of losing what we stand for. It is taking longer to get the money to where it is needed and our costs are going up. We are in danger of failing on one of our key objectives: to minimise the proportion of our turnover that is spent on administration. At the same time, we always need to be aware of the risk of bad publicity through making the wrong decisions. If we don't check applications thoroughly, funds may go to the 'wrong' place and if the newspapers get hold of the story we would run a real risk of losing the goodwill, and therefore the funds, from our many supporters*.'

Susan holds regular meetings with key stakeholders. One charity that handled a large number of applications for people in Nigeria told her of frequent complaints about the delays over the processing of the applications. A second charity representative complained that when he telephoned to find out the status of an application, the ARAPU staff did not seem to know where it was or how long it might be before it was complete. Furthermore, he felt that this lack of information was eroding his relationship with his own clients, some of whom were losing faith in him as a result, '*trust is so important in the relationship*', he explained.

Some of Susan's colleagues, while broadly agreeing with her anxieties over the organisation's responsiveness and efficiency, take a slightly different perspective. '*One of the really good things about Action Response is that we are more flexible than most charities. If there is a need, and if they need support until one of the larger charities can step in, then we will always consider a request for aid. I would not like to see any move towards high process efficiency harming our ability to be open-minded and consider requests that might seem a little unusual at first*.' (Jacqueline Horton, Applications Assessor)

Others see the charity as performing an important counselling role. '*Remember that we have gained a lot of experience in this kind of short-term aid. We are often the first people that are in a position to advise on how to apply for larger and longer-term funding. If we developed this aspect of our work we would again be fulfilling a need that is not adequately supplied at the moment*.' (Stephen Nyquist, Applications Assessor)

The Action Response Applications Processing Unit (ARAPU)

Potential aid recipients, or the intermediary charities representing them, apply for funds using a standard form. These forms can be downloaded from the internet or requested via a special helpline. Sometimes the application will come directly from an individual community leader but more usually it will come via an intermediary charity that can help the applicant to complete the form. The application is sent to ARAPU, usually by fax or post (some are submitted online, but few communities have this facility).

ARAPU employs seven applications assessors with support staff who are responsible for data entry, coding, filing and 'completing' (staff who prepare payment, or explain why no aid can be given). In addition, a board of non-paid trustees meets every Thursday, to approve the assessors' decisions. The unit's IT system maintains records of all transactions, providing an update on the number of applications received, approved, declined and payments allocated. These reports identify that the unit receives about 300 new applications per week and responds to about the same number (the unit operates a 35-hour week). But while the unit's financial targets are being met, the trend indicates that the cost per application is increasing. The target for the turnaround of an application, from receipt of application to response, is 20 days, and although this is not measured formally, it is generally assumed that turnaround time is longer than this. Accuracy has never been an issue, as all files are thoroughly assessed to ensure that all the relevant data are collected before the applications are processed. Productivity seems high and there is always plenty of work waiting for processing at each section, with the exception that the 'completers' are sometimes waiting for work to come from the committee on a Thursday. Susan has conducted an inspection of all sections' in-trays that has revealed a rather shocking total of about 2,000 files waiting within the process, not counting those waiting for further information.

Processing applications

The processing of applications is a lengthy procedure requiring careful examination by applications assessors, trained to make well-founded assessments in line with the charity's guidelines and values. Incoming applications are opened by one of the four 'receipt' clerks who check that all the necessary forms have been included in the application; the receipt clerks take about 10 minutes per application. These are then sent to the coding staff in batches, twice a day. The five coding clerks allocate a unique identifier to each application and key the information on the application into the system. The coding stage takes about 20 minutes for each application. Files are then sent to the senior applications assessors' secretary's desk. As an assessor becomes available, the secretary provides the next job in the line to the assessor.

About 100 of the cases seen by the assessors each week are put aside after only 10 minutes 'scanning' because the information is ambiguous, so further information is needed. The assessor returns these files to the secretaries, who write to the applicant (usually via the intermediate charity) requesting additional information, and return the file to the 'receipt' clerks who 'store' the file until the further information eventually arrives (usually between one and eight weeks). When it does arrive, the file enters the process and progresses through the same stages again. Of the applications that require no further information, around half (150) are accepted and half (150) declined. On average, those applications that are not 'recycled' take around 60 minutes to assess.

All the applications, whether approved or declined, are stored prior to ratification. Every Thursday the Committee of Trustees meets to formally approve the applications assessors' decisions. The committee's role is to sample the decisions to ensure that the guidelines of the charity are upheld. In addition, they will review any particularly unusual cases highlighted by the applications assessors. Once approved by the committee the files are then taken to the completion officers. There are three 'declines' officers whose main responsibility is to compile a suitable response to the applicant, pointing out why the application failed and offering, if possible, helpful advice. An experienced declines officer takes about 30 minutes to finalise the file and write a suitable letter. Successful files are passed to the four 'payment' officers where again the file is completed, letters (mainly standard letters) are created and payment instructions are given to the bank. This usually takes around 50 minutes, including dealing with any queries from the bank about payment details. Finally, the paperwork itself is sent, with the rest of the file, to two 'dispatch' clerks who complete the documents and mail them to the applicant. The dispatch activity takes, on average, 10 minutes for each application.

The feeling among the staff is generally good. When Susan consulted the team, they said their work was clear and routine, but their life was made difficult by charities that rang expecting them to be able to tell them the status of an application they had submitted. It could take them hours, sometimes days, to find any individual file. Indeed, two of the 'receipt' clerks are now working almost full time on this activity. They also said that charities frequently complained that decision-making seemed slow.

Tinpixels/E+/Getty Images

Questions

1. What objectives should the ARAPU process be trying to achieve?
2. What is the main problem with the current ARAPU process?'
3. How could the ARAPU process be improved?

Applying the principles

Some of these exercises can be answered by reading the chapter. Others will require some general knowledge of business activity and some might require an element of investigation. **Model answers for the first two questions are available on the student companion website**. *For model answers to the other questions in this section, please ask your tutor/lecturer.*

1. Most of us are familiar with 'drive-through' fast-food operations. Think about (or better still, visit) a drive-through service and try mapping what you can see (or remember) of the process (plus what you can infer from what may be happening 'behind the scenes').

2. '*It is a real problem for us*', said Angnyeta Larson. '*We now have only ten working days between all the expense claims coming from the departmental coordinators and authorising payments on the next month's payroll. This really is not long enough and we are already having problems during peak times*.' Angnyeta was the department head of the internal financial control department of a metropolitan authority in southern Sweden. Part of her department's responsibilities included checking and processing expense claims from staff throughout the metropolitan authority and authorising payment to the salaries section. She had 12 staff who were trained to check expense claims and all of them were devoted full time to processing the claims in the two weeks (10 working days) prior to the deadline for informing the salaries section. The number of claims submitted over the year averaged around 3,200, but this could vary between 1,000 during the quiet summer months up to 4,300 in peak months. Processing claims involved checking receipts, checking that claims met with the strict financial allowances for different types of expenditure, checking all calculations, obtaining more data from the claimant if necessary, and (eventually) sending an approval notification to salaries. The total processing time took on average 20 minutes per claim.

   a) How many staff does the process need on average, for the lowest demand and for the highest demand?
   b) If a more automated process involving electronic submission of claims could reduce the average processing time to 15 minutes, what effect would this have on the required staffing levels?
   c) If department coordinators could be persuaded to submit their batched claims earlier (not always possible for all departments) so that the average time between submission of the claims to the finance department and the deadline for informing the salaries section was increased to 15 working days, what effect would this have?

3. The headquarters of a major creative agency offered a service to all its global subsidiaries that included the preparation of a budget estimate that was submitted to potential clients when making a 'pitch' for new work. This service had been offered previously only to a few of the group's subsidiary companies. Now that it was to be offered worldwide, it was deemed appropriate to organise the process of compiling budget estimates on a more systematic basis. It was estimated that the worldwide demand for this service would be around 20 budget estimates per week and that, on average, the staff who would put together these estimates would be working a 35-hour week. The elements within the total task of compiling a budget estimate are shown in Table 6.4.

   a) What is the required cycle time for this process?
   b) How many people will the process require to meet the anticipated demand of 20 estimates per week?
   c) Assuming that the process is to be designed on a 'long–thin' basis, what elements would each stage be responsible for completing? And what would be the balancing loss for this process?
   d) Assuming that instead of the long–thin design, two parallel processes are to be designed, each with half the number of stations of the long–thin design, what now would be the balancing loss?

Table 6.4 The elements within the total task of compiling a budget estimate

Element	*Time (mins)*	*What element(s) must be done prior to this one?*
A – obtain time estimate from creatives	20	None
B – obtain account handler's deadlines	15	None
C – obtain production artwork estimate	80	None
D – make preliminary budget calculations	65	A, B and C
E – check on client budget	20	D
F – check on resource availability and adjust estimate	80	D
G – complete final budget estimate	80	E and F

4. At the theatre, the interval during a performance of *King Lear* lasts for 20 minutes and in that time 86 people need to use the toilet cubicles. On average, a person spends 3 minutes in the cubicle. There are 10 cubicles available.

a) Does the theatre have enough toilets to deal with the demand?

b) If there are not enough cubicles, how long should the interval be to cope with demand?

5. A gourmet burger shop has a daily demand for 250 burgers and operates for 10 hours.

a) What is the required cycle time in minutes?

b) Assuming that each burger takes 7.2 minutes of work, how many servers are required?

c) If the burger shop has a three-stage process for making burgers, and stage 1 takes 2.0 minutes, stage 2 takes 3.0 minutes and stage 3 takes 2.2 minutes, what is the balancing loss for the process?

Notes on chapter

1 The information on which this example is based is taken from: Zhang, S. (2016) 'How to fit the world's biggest indoor waterfall in an airport', *Wired*, 9 July; Airport Technology (2014) 'Terminal 4, Changi International Airport', airporttechnology.com; Driver, C. (2014) 'And the winners are . . . Singapore crowned the best airport in the world (and Heathrow scoops top terminal)', *Mail*online, 28 March, https://www.dailymail.co.uk/travel/article-2591405/Singapore-crowned-best-airport-world-Heathrow-scoops-terminal.html [accessed 17 September 2020].

2 The information on which this example is based is taken from: Qureshi, W. (2020) 'Ecover relaunches biodegradable detergents in PCR plastic', *Packaging News*, 21 January; Cornwall, S. (2013) 'Ecover announces world-first in plastic packaging', *Packaging Gazette*, 7 March; Ecover website, www.ecover.com [accessed 17 September 2020].

3 Fulton Suri, J. (2003) 'The experience of evolution: Developments in design practice', *The Design Journal*, 6 (2), pp. 39–48; Forlizzi, J. and Ford, S. (2000 'The building blocks of experience: An early framework for interaction designers', *Designing Interactive Systems 2000*, ACM Press, pp. 419–423.

4 Shostack, G.L. (1984) 'Designing services that deliver', *Harvard Business Review*, 62 (1), pp. 133–139.

5 The information on which this example is based is taken from: Matthews, T. and Trim, L. (2019) 'London Underground: Why it would be better if we stood on both sides of the escalators', *MyLondon* Local News, 19 August; Sleigh, S. (2017) 'TfL scraps standing only escalators – despite trial being deemed a "success"', *Evening Standard*, 8 March.

Taking it further

Curedale, R. ***(2018) Service Design Process & Methods, 3rd edition, Design Community College Inc.*** *One of the relatively few books that deals with service design specifically*.

Damelio, R. ***(2011) The Basics of Process Mapping, 2nd edition, Productivity Press.*** *A practitioner book that is both very comprehensive and up to date*.

Hammer, M. ***(1990) 'Reengineering work: Don't automate, obliterate', Harvard Business Review, July–August.*** *This is the paper that launched the whole idea of business processes and process management in general to a wider managerial audience. Slightly dated but worth reading*.

Harrington, H.J. ***(2011) Streamlined Process Improvement, McGraw Hill Professional.*** *Practical and insightful*.

Holweg, M., Davies, J., De Meyer, A., Lawson, B. and Schmenner, R.W. ***(2018) Process Theory: The principles of operations management, Oxford University Press.*** *Don't be put off by the title, it is not particularly theoretical*.

Hopp, W.J. and Spearman, M.L. ***(2011) Factory Physics, 3rd edition, Waveland Press.*** *Very technical so don't bother with it if you aren't prepared to get into the maths. However, some fascinating analysis, especially concerning Little's Law*.

Page, S. ***(2015) The Power of Business Process Improvement: 10 simple steps to increase effectiveness, efficiency, and adaptability, 2nd edition, AMACOM.*** *Provides a practical guide to process design*.

Ramaswamy, R. ***(1996) Design and Management of Service Processes: Keeping customers for life, Addison-Wesley Longman.*** *A relatively technical approach to process design in a service environment*.

Slack, N. ***(2017) The Operations Advantage: A practical guide to making operations work, Kogan Page.*** *The chapter on 'internal processes' expands on some of the issues discussed here*.

Sparks, W. ***(2016) Process Mapping Road Trip: Improve organizational workflow in five steps, Promptitude Publishing.*** *A practitioner's guide – straightforward and sensible*.

Supplement to Chapter 6
Queuing analysis

Queuing analysis (in many parts of the world it is called 'waiting line analysis') is often explained purely in terms of customers being processed through service operations. This is misleading. Although queuing analysis can be particularly important in service operations, especially where customers really do 'queue' for service, the approach is useful in any kind of operation. Figure 6.17 shows the general form of queuing analysis. Customers arrive according to some probability distribution and wait to be processed (unless part of the operation is idle); when they have reached the front of the queue, they are processed by one of the *n* parallel 'servers' (their processing time also being described by a probability distribution), after which they leave the operation. There are many examples of this kind of system and Table 6.5 illustrates some of these. All of these examples can be described by a common set of elements that define their queuing behaviour:

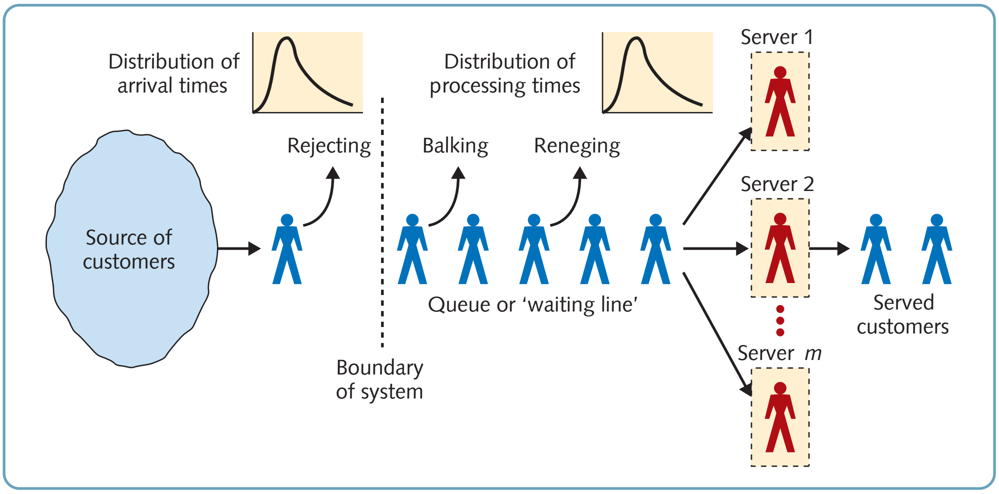


Figure 6.17 **The general form of queuing analysis**

Table 6.5 Examples of processes that can be analysed using queuing analysis

Operation	*Arrivals*	*Processing capacity*
Bank	Customers	Tellers
Supermarket	Shoppers	Checkout staff
Hospital clinic	Patients	Doctors
Graphic artist	Commissions	Artists
Custom cake decorators	Orders	Cake decorators
Ambulance service	Emergencies	Ambulances with crews
Telephone switchboard	Calls	Telephonists
Maintenance department	Breakdowns	Maintenance staff

The source of customers, sometimes called the calling population, is the source of supply of customers. In queue management 'customers' are not always human. 'Customers' could, for example, be trucks arriving at a weighbridge, orders arriving to be processed or machines waiting to be serviced, etc.

The arrival rate is the rate at which customers needing to be served arrive at the server or servers. Rarely do customers arrive at a steady and predictable rate. Usually, there is variability in their arrival rate. Because of this it is necessary to describe arrival rates in terms of probability distributions.

The queue, or waiting line, is formed by the customers waiting to be served. If there is relatively little limit on how many customers can queue at any time, we can assume that, for all practical purposes, an infinite queue is possible. Sometimes, however, there is a limit to how many customers can be in the queue at any one time.

Queue discipline is the set of rules that determine the order in which customers waiting in the queue are served. Most simple queues, such as those in a shop, use a first come, first served queue discipline.

Servers are the facility that processes the customers in the queue. In any queuing system there may be any number of servers configured in different ways. In Figure 6.17 servers are configured in parallel, but some systems may have servers in a series arrangement. There is also likely to be variation in how long it takes to process each customer. Therefore, processing time, like arrival time, is usually described by a probability distribution.

Calculating queue behaviour

Management scientists have developed formulae that can predict the steady-state behaviour of different types of queuing system. Unfortunately, many of these formulae are extremely complicated, especially for complex queuing systems, and are beyond the scope of this text. In fact, in practice, computer programs are almost always used to predict the behaviour of queuing systems. However, studying queuing formulae can illustrate some useful characteristics of the way queuing systems behave. Moreover, for relatively simple systems, using the formulae (even with some simplifying assumptions) can provide a useful approximation to process performance.

Notation

There are several different conventions for the notation used for different aspects of queuing system behaviour. It is always advisable to check the notation used by different authors before using their formulae. We shall use the following notation:

t_a = average time between arrivals

r_a = arrival rate (items per unit time) $= 1/t_a$

c_a = coefficient of variation of arrival times

m = number of parallel servers at a station

t_e = mean processing time

r_e = processing rate (items per unit time) $= m/t_e$

c_e = coefficient of variation of process time

u = utilisation of station $= r_a/r_e = (r_a t_e)/m$

WIP = average work-in-process (number of items) in the queue

WIP_q = expected work-in-process (number of items) in the queue

W_q = expected waiting time in the queue

W = expected waiting time in the system (queue time + processing time)

Variability

The concept of variability is central to understanding the behaviour of queues. If there were no variability there would be no need for queues to occur because the capacity of a process could be relatively easily adjusted to match demand. For example, suppose one member of staff (a server) serves customers at a bank counter who always arrive exactly every 5 minutes (i.e. 12 per hour). Also suppose that every customer takes exactly 5 minutes to be served, then because:

a) the arrival rate is the processing rate and

b) there is no variation,

no customer need ever wait because the next customer will arrive when, or after, the previous customer leaves. That is, $WIP_q = 0$.

Also, in this case, the server is working all the time, again because exactly as one customer leaves, the next one is arriving. That is, $u = 1$.

Even with more than one server, the same may apply. For example, if the arrival time at the counter is 5 minutes (12 per hour) and the processing time for each customer is now always exactly 10 minutes, the counter would need two servers, and because:

a) arrival rate is processing rate m, and

b) there is no variation,

then again, $WIP_q = 0$, and $u = 1$.

Of course, it is convenient (but unusual) if arrival rate/processing rate = a whole number. When this is not the case (as in this simple example with no variation):

$$
\begin{aligned}
\text{utilisation} &= \text{processing rate/(arrival rate } m) \\
\text{for example, if arrival rate, } r_a &= 5 \text{ minutes} \\
\text{processing rate, } r_e &= 8 \text{ minutes} \\
\text{number of servers, } m &= 2 \\
\text{then, utilisation, } u &= 8/(5 \times 2) = 0.8 \text{ or } 80\%.
\end{aligned}
$$

Incorporating variability

The previous examples were not realistic because they assumed no variation in arrival or processing times. We also need to take into account the variation around these means. To do that, we need to use a probability distribution. Figure 6.18 contrasts two processes with different arrival distributions. The units arriving are shown as people, but they could be jobs arriving at a machine, trucks needing servicing or any other uncertain event. The top example shows low variation in arrival time where customers arrive in a relatively predictable manner. The bottom example has the same average number of customers arriving but this time they arrive unpredictably with sometimes long gaps between arrivals and at other times two or three customers arriving close together. We could do a similar analysis to describe processing times.

In Figure 6.18 high arrival variation has a distribution with a wider spread (called 'dispersion') than the distribution describing lower variability. Statistically the usual measure for indicating the spread of a distribution is its standard deviation α. But variation does not only depend on standard deviation. For example, a distribution of arrival times may have a standard deviation of 2 minutes. This could indicate very little variation when the average arrival time is 60 minutes. But it would mean a very high degree of variation when the average arrival time is 3 minutes. Therefore, to normalise standard deviation, it is divided by the mean of its distribution. This measure is called the coefficient of variation of the distribution. So,

$$
\begin{aligned}
c_a &= \text{coefficient of variation of arrival times} = \sigma_a/t_a \\
c_e &= \text{coefficient of variation of processing times} = \sigma_e/t_e
\end{aligned}
$$

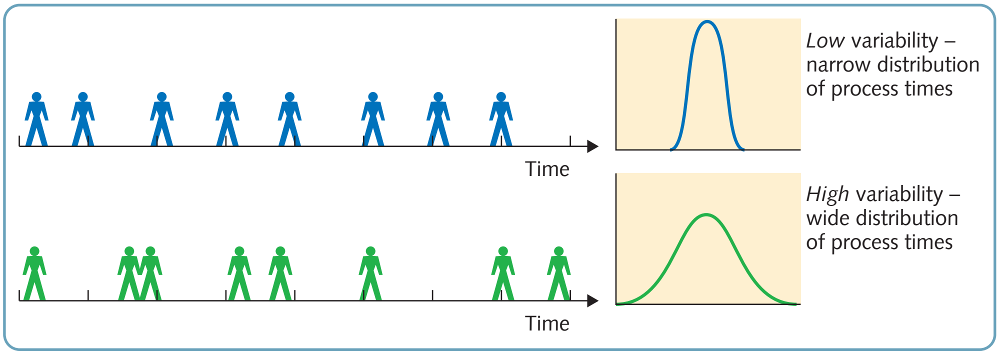


Figure 6.18 Low and high arrival variation

Incorporating Little's Law

Little's Law (introduced earlier in the chapter) describes the relationship between the cycle time, the work-in-process and the throughput time of the process. It is denoted by the following simple relationship:

$$\text{Throughput time} = \text{Work-in-process} \times \text{cycle time}$$

So

$$\text{Work-in-process} = \text{Throughput time/cycle time}$$

Or

$$\text{WIP} = T/C$$

And

$$\text{Cycle time} = 1/\text{arrival rate}$$

Little's Law can help to understand queuing behaviour. Consider the queue in front of a station:

Work-in-process in the queue = arrival rate at the queue (equivalent to 1/cycle time) × waiting time in the queue (equivalent to throughput time)

$$\text{WIP}_q = r_a \times W_q$$

and

waiting time in the whole system = the waiting time in the queue + the average process time at the station

$$W = W_q + t_e$$

We will use this relationship later to investigate queuing behaviour.

Types of queuing system

Conventionally queuing systems are characterised by four parameters:

A = the distribution of arrival times (or more properly, inter-arrival times, the elapsed times between arrivals)
B = the distribution of process times
m = the number of servers at each station
b = the maximum number of items allowed in the system.

The most common distributions used to describe A or B are either:

i) the exponential (or Markovian) distribution denoted by M
ii) the general (for example normal) distribution denoted by G.

So, for example, an M/G/1/5 queuing system would indicate a system with exponentially distributed arrivals, process times described by a general distribution such as a normal distribution, with one server and a maximum number of five items allowed in the system. This type of notation is called Kendall's notation.

Queuing theory can help us investigate any type of queuing system, but in order to simplify the mathematics, we shall here deal only with the two most common situations. Namely: M/M/m – the exponential arrival and processing times with *m* servers and no maximum limit to the queue; and G/G/m – general arrival and processing distributions with *m* servers and no limit to the queue.

And we will start by looking at the simple case when M = 1.

For M/M/1 queuing systems

The formulae for this type of system are as follows:

$$\text{WIP} = \frac{u}{1-u}$$

Using Little's Law:

$$\text{WIP} = \text{throughput time/cycle time}$$

$$\text{Throughput time} = \text{WIP cycle time} = \text{WIP/arrival rate}$$

Then:

$$\text{Throughput time} = \frac{u}{1-u} \times \frac{1}{r_a} = \frac{t_e}{1-u}$$

And since throughput time in the queue = total throughput time – average processing time:

$$\begin{aligned} W_q &= W - t_e \\ &= \frac{t_e}{1-u} - t_e \\ &= \frac{t_e - t_e(1-u)}{1-u} = \frac{t_e - t_e - ut_e}{1-u} \\ &= \frac{u}{(1-u)} t_e \end{aligned}$$

Again, using Little's Law:

$$\text{WIP}_q = r_a \times W_q = \frac{u}{(1-u)} t_e \, r_a$$

And since:

$$u = \frac{r_a}{r_e} = r_a \, t_e$$

$$r_a = \frac{u}{t_e}$$

Then:

$$\begin{aligned} \text{WIP}_q &= \frac{u}{(1-u)} \times t_e \times \frac{u}{t_e} \\ &= \frac{u^2}{(1-u)} \end{aligned}$$

For M/M/m systems

When there are m servers at a station the formula for waiting time in the queue (and therefore all other formulae) needs to be modified. Again, we will not derive these formulae but just state the modified one as:

$$W_q = \frac{u^{\sqrt{2(m+1)}-1}}{m(1-u)} t_e$$

from which the other formulae can be derived as before.

For G/G/1 systems

The assumption of exponential arrival and processing times is convenient as far as the mathematical derivations of various formulae are concerned. However, in practice, process times in particular are rarely truly exponential. This is why it is important to have some idea of how a G/G/1 and G/G/M queue behaves. However, as exact mathematical relationships are not possible with such distributions, therefore some kind of approximation is needed. The one here is in common use, and although it is not always accurate, it is accepted for practical purposes. For G/G/1 systems the formula for waiting time in the queue is as follows:

$$W_q = \left(\frac{c_a^2 + c_e^2}{2}\right)\left(\frac{u}{1-u}\right) t_e$$

There are two points to make about this equation. The first is that it is exactly the same as the equivalent equation for an M/M/1 system but with a factor to take account of the variability of the arrival and process times. The second is that this formula is sometimes known as the VUT formula because it describes the waiting time in a queue as a function of:

V = the variability in the queuing system
U = the utilisation of the queuing system (that is, demand versus capacity) and
T = the processing times at the station.

In other words, we can reach the intuitive conclusion that queuing time will increase as variability, utilisation or processing time increase.

For G/G/m systems

The same modification applies to queuing systems using general equations and m servers. The formula for waiting time in the queue is now as follows:

$$W_q = \left(\frac{c_a^2 + c_e^2}{2}\right)\left(\frac{u^{\sqrt{2(m+1)}-1}}{m(1-u)}\right) t_e$$

Worked example

'I can't understand it'

'I can't understand it. We have worked out our capacity figures and I am sure that one member of staff should be able to cope with the demand. We know that customers arrive at a rate of around 6 per hour and we also know that any trained member of staff can process them at a rate of 8 per hour. So why is the queue so large and the wait so long? Have a look at what is going on there please.'

Sarah knew that it was probably the variation, both in customers arriving and in how long it took each of them to be processed, that was causing the problem. Over a two-day period when she was told that demand was more or less normal, she timed the exact arrival times and processing times of every customer. Her results were as follows:

The coefficient of variation of customer arrivals, $c_a = 1$

The coefficient of variation of processing time, $c_e = 3.5$

The average arrival rate of customers, $r_a = 6$ per hour

Therefore the average inter-arrival time $= 10$ minutes

The average processing rate, $r_e = 8$ per hour

Therefore the average processing time $= 7.5$ minutes

Utilisation of the single server, $u = 6/8 = 0.75$

Using the waiting time formula for a G/G/1 queuing system:

$$W_q = \left(\frac{1 + 12.25}{2}\right)\left(\frac{0.75}{1 - 0.75}\right)7.5$$
$$= 6.625 \times 3 \times 7.5 = 149.06 \text{ mins}$$
$$= 2.48 \text{ hours}$$

Also, because:

$$\text{WIP}_q = \text{cycle time} \times \text{throughput time}$$
$$\text{WIPq} = 6 \times 2.48 = 14.68\%$$

So, Sarah had found out that the average wait that customers could expect was 2.48 hours and that there would be an average of 14.68 people in the queue.

'OK, so I see that it's the very high variation in the processing time that is causing the queue to build up. How about investing in a new computer system that would standardise processing time to a greater degree? I have been talking with our technical people and they reckon that, if we invested in a new system, we could cut the coefficient of variation of processing time down to 1.5. What kind of a difference would this make?'

Under these conditions with $c_e = 1.5$

$$W_q = \left(\frac{1 + 12.25}{2}\right)\left(\frac{0.75}{1 - 0.75}\right)7.5$$
$$= 1.625 \times 3 \times 7.5 = 36.56 \text{ mins}$$
$$= 0.61 \text{ hours}$$

Therefore:

$$\text{WIP}_q = 6 \times 0.61 = 3.66$$

In other words, reducing the variation of the process time has reduced average queuing time from 2.48 hours down to 0.61 hours and has reduced the expected number of people in the queue from 14.68 down to 3.66.

Worked example

Big Bad Bank

Big Bad Bank wishes to decide how many staff to schedule during its lunch period. During this period, customers arrive at a rate of 9 per hour and the enquiries that customers have (such as opening new accounts, arranging loans, etc.) take, on average, 15 minutes to deal with. The bank manager feels that four staff should be on duty during this period but wants to make sure that the customers do not wait more than 3 minutes on average before they are served. The manager has been told by his small daughter that the distributions that describe both arrival and processing times are likely to be exponential. Therefore:

$r_a = 9$ per hour, therefore
$t_a = 6.67$ minutes
$r_e = 4$ per hour, therefore
$t_e = 15$ minutes

The proposed number of servers, $m = 4$ therefore the utilisation of the system, $u = 9/4 \times 4 = 0.5625$.

From the formula for waiting time for an M/M/m system:

$$W_q = \frac{u^{\sqrt{2(m+1)}-1}}{m(1 - u)} t_e$$

$$W_q = \frac{u^{\sqrt{2(m+1)}-1}}{m(1-u)} t_e$$

$$W_q = \frac{0.5625^{\sqrt{10}-1}}{4(1-0.5625)} \times 0.25$$

$$= \frac{0.5625^{2.162} - 1}{1.75} \times 0.25$$

$$= 0.042 \text{ hours}$$

$$= 2.52 \text{ minuts}$$

$$W_q = \frac{0.5625^{\sqrt{10}-1}}{4(1-0.5625)} \times$$

$$= \frac{0.5625^{2.162} - 1}{1.75} \times 0.25$$

$$= 0.042 \text{ hours}$$

$$= 2.52 \text{ minutes}$$

Therefore the average waiting time with four servers would be 2.52 minutes, which is well within the manager's acceptable waiting tolerance.

Taking it further

Hopp, W. J. and Spearman, M. L. ***(2011) Factory Physics, 3rd edition, Waveland Press.*** *Very technical so don't bother with it if you aren't prepared to get into the maths. However, some fascinating analysis, especially concerning Little's Law*.

7 Supply chain management

Introduction

Given the typically large proportion of activities that are outsourced by operations, managing supply networks is a particularly vital activity. In Chapter 4 we explored the structure and scope of operations and the network of which they are part; by contrast, this chapter is more concerned with how supply chains and networks are managed on an ongoing basis. This involves determining key performance objectives for the supply network, deciding on supplier relationships (transactional versus partnership), developing sourcing strategies for different products and services, selecting appropriate suppliers and negotiating the terms of their engagement, managing day-to-day supply, improving suppliers' capabilities over time and attempting to mitigate supply chain dynamics. These activities are usually the responsibility of the purchasing or procurement function within the business. Purchasing should provide a vital link between the operation itself and its suppliers. It should understand the requirements of all the processes within its own operation and also the capabilities of the suppliers who could potentially provide services for the operation. Figure 7.1 shows the position of the ideas described in this chapter in the general model of operations management.

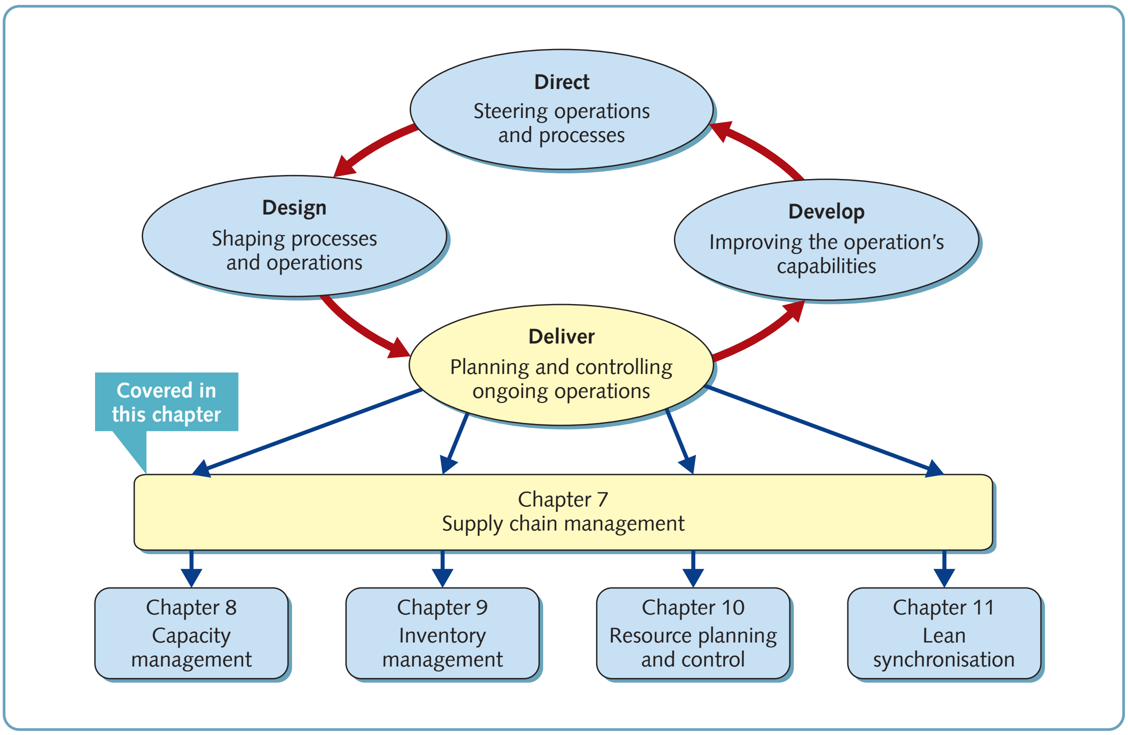


Figure 7.1 Supply chain management is the management of the relationships and flows of products and services between operations and processes

EXECUTIVE SUMMARY

7.1 Is the nature of supply chain management understood?

↓

7.2 What type of supply chain relationship should be adopted?

↓

7.3 Has the sourcing configuration been determined?

↓

7.4 Has the approach to supplier selection and negotiation been determined?

↓

7.5 Are supply chains run effectively?

↓

7.6 Are supply chain dynamics under control?

7.1 Is the nature of supply chain management understood?

Supply chain management (SCM) is concerned with managing the flows of information, products and services between the many 'strings' (or chains) of operations and processes that form an organisation's supply network. It is an essential activity because the overall success of any individual firm is fundamentally impacted by the effectiveness of its supply network. In addition, most principles applied to *external* (organisation-to-organisation) networks can also be applied to *internal* (process-to-process) networks. The fundamental objective of SCM is to satisfy the needs of the end customer. So, each operation (or process) in the network should contribute to whatever mix of quality, speed, dependability, flexibility, cost and sustainability that the end customer requires. In establishing objectives for the supply network, the issue of alignment is important for managers to understand. Typically, efficient (or lean) supply networks are suitable for stable 'functional' services and products, while responsive (or agile) supply networks are more suitable for less predictable 'innovative' services and products.

7.2 What type of supply chain relationship should be adopted?

Supply chain relationships can be categorised by *what* is outsourced (from everything to nothing) and *who* supplies the products or services. The decision concerning *what* to outsource is addressed in Chapter 4 on operations scope and structure. In considering *who* supplies in this chapter, we examine market-based 'transactional' (or contractual) relationships at one end of the spectrum and longer-term 'partnership' (or collaborative) relationships at the other. Each have their advantages and disadvantages, though sometimes the type of relationship adopted may be dictated by the structure of the market itself. If the number of potential suppliers is small, there are few opportunities to use market-based mechanisms to gain leverage.

7.3 Has the sourcing configuration been determined?

Once managers have determined the type of supply chain relationship that is appropriate, they must consider preferences in relation to sourcing configuration. This involves deciding on the alternative sourcing methods (*multiple sourcing, single* *sourcing, delegated sourcing* or *parallel sourcing*) that should be used for different categories of services or products. These categories typically relate to different levels of supply market risk and criticality of the purchase to the business, and form four quadrants – *non-critical, bottleneck, leverage* and *strategic*.

7.4 Has the approach to supplier selection and negotiation been determined?

There are a number of steps in engaging suppliers. Initial qualification aims to establish a minimum threshold for a supplier to be considered viable and often relates to *technical capability*, *operations capability* and *financial capability*. Measurement criteria should then be determined, factoring in the relative importance of different performance objectives. This is followed by information collection to inform the supplier selection decision. When making the final selection, decision-makers must trade-off supplier attributes, often using scoring assessment methods. Once the preferred supplier has been identified, negotiation takes place to (try to) agree the key terms of engagement. For those involved in this activity, it is important to understand the various tactics (including *emotion*, *logic*, *threats*, *bargaining* and *compromise*) that may be employed during the negotiation process.

7.5 Are supply chains run effectively?

There are a number of important issues to consider when attempting to effectively run the many supply chains that form a supply network. First, how to manage the two-way flow of information between customers and suppliers; and factoring in the multi-layered nature of such relationships. Service-level agreements are often employed to help formalise expectations between parties in a supply network. Second, how to minimise the common mismatches between customer and supplier perceptions of both requirements and performance. Third, how to develop suppliers, especially those in partnership relationships. Fourth, how to effectively manage logistics (or distribution), which is particularly important for those in more product-oriented settings. The 'Internet of Things' (IoT) has had a significant impact on logistics in providing firms with much more accurate information on where products are in the supply network.

7.6 Are supply chain dynamics under control?

Supply chains have a dynamic of their own that is often called the *bullwhip effect*, which has various detrimental financial and operational consequences. It means that relatively small changes at the demand end of the chain increasingly amplify into large disturbances as they move upstream. Four methods can be used to reduce this effect. First, channel alignment through standardised planning and control methods allows for easier coordination of the whole chain. Second, improving the operational efficiency of each part of the chain prevents local errors multiplying to affect the whole chain. Third, improved forecasts allow for reduced inventory holding requirements and less excess capacity to guarantee service levels. Fourth, information sharing can prevent overreaction to immediate stimuli and give a better view of the whole supply network. Blockchain-enabled systems represent an important development in attempts by organisations to improve the transparency and fidelity of information within their supply networks.

7.1 Diagnostic question: *Is the nature of supply chain management understood?*

Operations do not exist in isolation. No matter how good an individual operation is, its overall performance is partly a function of the entire supply network of which it is a part. Just consider the impact on supply caused by the disruption to supply chains from factory fires, tsunamis, pandemics, trade disputes and so on. It can be a demonstration that it is the whole network that determines the quality, speed, dependability, flexibility, cost and sustainability of the final services or products that are delivered to end customers. In this section, we explain what supply chain management is and why it matters so much for firms. We then examine the key objectives that should be considered by decision-makers looking to improve supply chain performance.

Managing supply chains and supply networks

Supply chain management (SCM) is the management of the relationships and flows between the 'string' (or chain) of operations and processes that deliver value in the form of products and services to the ultimate consumer. It is a holistic approach to managing across the boundaries of companies and of processes. As illustrated in Figure 7.2, supply *chains* are technically different to supply *networks*. A supply network is *all* the operations that are linked together so as to

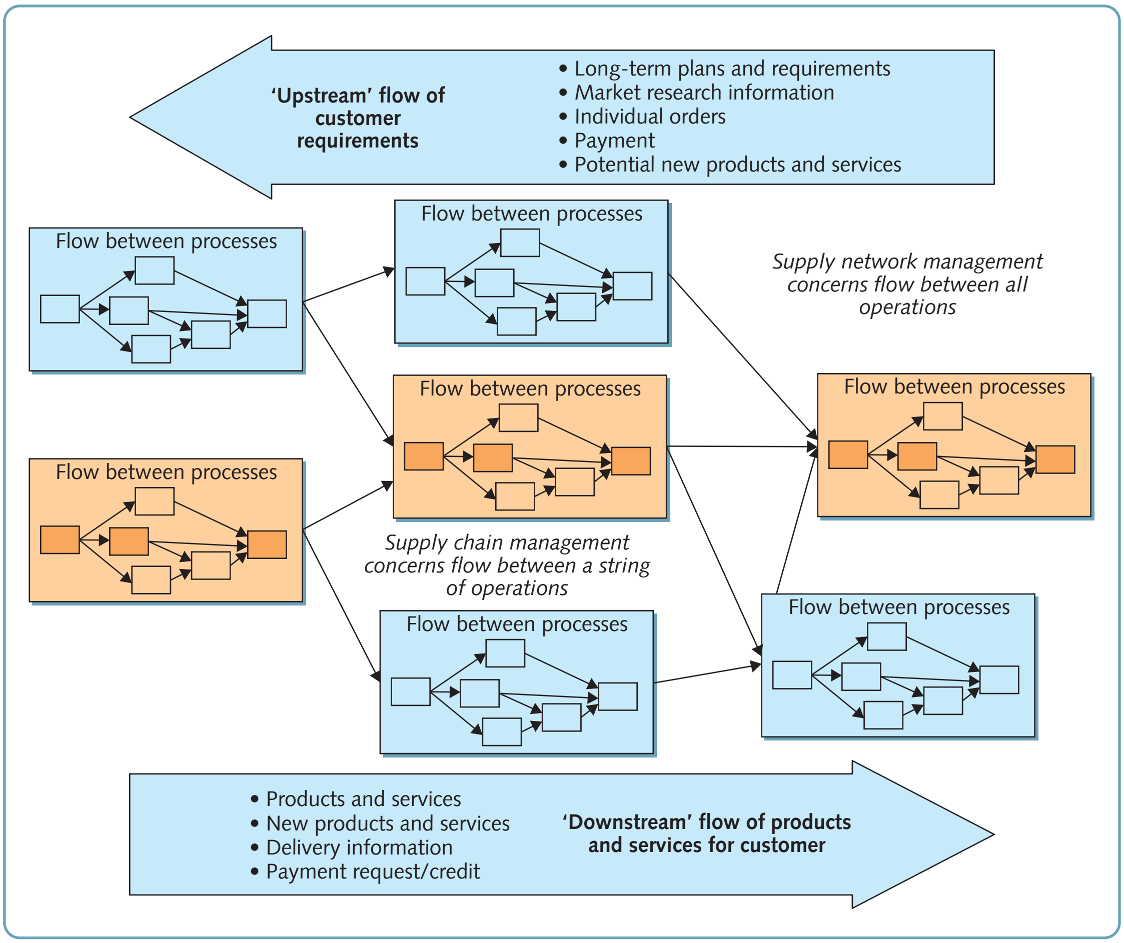


Figure 7.2 Managing supply chains and supply networks is equally concerned with flows of information as with flows of products or services

provide products and services through to end customers. In large supply networks there can be many hundreds of supply chains of linked operations passing through a single operation. Confusingly, the terms supply network and supply chain are often (mistakenly) used interchangeably. It is also worth noting that the 'flows' in supply chains are not restricted to the downstream flow of products and services from suppliers through to customers. Although the most obvious failures in supply chain management occur when downstream product or service flows fail to meet customer requirements, the root causes may be failures in the upstream flows of information. As such, supply chain management is as much concerned with managing information flows (upstream and downstream) as it is with managing the flow of products and services.

Supply chain management applies to non-physical flow

The vast majority of books, blogs and articles that focus on the challenges of supply chain management, continue to focus on 'material transformation' operations – operations that are concerned with the creation, movement, storage or sale of physical products. However, supply chain management applies equally to operations with largely or exclusively intangible inputs and outputs, such as financial services, retail shopping malls, insurance providers, healthcare operations, consultants, universities and so on. All these operations have suppliers and customers, they all purchase services, they all have to choose how they get their services to consumers. In other words, they all have to manage their supply chains. In reality, all supply networks, even ones that transform physical items, have service elements – again referring back to Chapter 1, most operations supply a mixture of products *and* services.

> OPERATIONS PRINCIPLE
> ***Supply chain management applies equally to non-physical flows between operations and processes as well as physical flows.***

Internal and external supply chains

Although we often describe supply chains as an interconnection of 'organisations', this does not necessarily mean that these 'organisations' are distinctly separate entities belonging to and managed by different owners. In Chapter 1, we pointed out how the idea of networks can be applied, not just at the supply network level of 'organisation-to-organisation' relationships, but also at the 'process-to-process' within-operation level and even at the 'resource-to-resource' process level. We also introduced the idea of internal customers and suppliers. Put these two related ideas together and one can understand how many of the issues that we will be discussing in the context of 'organisation-to-organisation' supply networks can also provide insight for internal 'process-to-process' supply networks.

> OPERATIONS PRINCIPLE
> ***Supply chain management concepts apply to internal relationships between processes as well as the external relationships between operations.***

Case example

Zipline's drone-enabled supply network[1]

When drones started to be deployed in supply chain management, one of their most successful users was not a big logistics provider but Zipline, who have built the world's largest autonomous logistics network, operating at national scale across multiple territories, delivering blood and medical supplies in parts of the developing world. In Rwanda, Zipline helped establish a valuable alternative to road transportation for high-value, low-weight items. Although the country was investing in its transport links, over 80 per cent of its road network remained unsurfaced, making road travel challenging at best, and almost impossible during the country's rainy season. For critical medical supplies, where delivery delays meant significant negative patient outcomes, the problem was acute. Also, medical supplies were hard to forecast or had short shelf lives, ideal for drone delivery, as were orders that dipped into safety stocks of drugs. With a small fleet of autonomous drones, each with carrying capacity of 1.5 kg (equivalent to 3×500 ml of blood) and capable of delivering anywhere within a 22,500 km^2 service area, Zipline was able to establish a highly effective delivery network for medical supplies, capable of reaching over half of Rwanda's 12 million population.

▶

Given the urgency of many of its orders, the firm focused on reducing the time to get drones airborne. One innovation was moving the GPS circuitry from the drone to its battery, enabling continuous GPS connection and saving precious minutes by establishing a stable signal prior to launch. Modular construction enabled orders to be placed inside the fuselage, which was put on the launcher, after which its wing section and battery module were attached. This made the drone easier to handle and ensured that problems picked up in pre-flight checks could be resolved quickly by switching out a faulty module. The often lengthy pre-flight checking was reduced by using a mobile inspection application and computer vision algorithms. The overall effect was to reduce the time between order receipt and drone launch to just five minutes. After which, a motorised launcher accelerated the drones up to their maximum speed of 100 km per hour in just 0.3 seconds. Not only was speed important, so was *dependability* of service. Having a back-up for each of the critical components on the drone, in case the primary component failed, helped reduce failed deliveries. At its final destination, the drone would simply drop its medical supplies box using a parachute made from wax paper and biodegradable tape, removing the need for delivery site infrastructure.

For some clients, Zipline operated a 'cross-docking' solution. Here, a client would prepare its packages which would then be consolidated and sent to Zipline's distribution centres. The firm would then schedule delivery with end recipients at a time that was convenient for them. Other clients preferred to use Zipline as a third-party logistics (3PL) provider. This involved the firm receiving inventory from its clients and holding them in their distribution centres ready to pack and ship rapidly as soon as orders were received. Ultimately, Zipline's aim was to create a supply network that was both efficient (lean) when it could be, but responsive (agile) when it needed to be. Higher costs of delivery using drones were inevitable, but these were partially offset against savings from reduced inventory in the supply network and less product obsolescence. In addition, the benefits in terms of speed, dependability and, most critically, patient health were substantial.

RUTH MCDOWALL/AFP/Getty Images

Supply chain objectives

The objectives of supply chain management are similar to those for individual operations – to deliver services and products to end customers that meet their expectations in terms of quality, speed, dependability, flexibility, cost and sustainability.

Quality

The quality of a product or service when it reaches the customer is a function of the quality performance of every operation in the chain that supplied it. The implication of this is that errors in each stage of the chain can multiply in their effect on end-customer service. For example, if each of seven stages in a supply chain has a 1 per cent error rate, only 93.2 per cent of products or services will be of good quality on reaching the end customer (0.99^7). This is why only by every stage taking some responsibility for its own *and its suppliers'* performance can a supply chain achieve high end-customer quality.

Speed

Speed has two meanings in a supply chain context. The first is how fast customers can be served (the elapsed time between a customer requesting a product or service and receiving it in full). However, fast customer response can be achieved simply by over-resourcing or overstocking within the supply chain. For example, very large stocks in a retail operation can reduce the chances of stock-out to almost zero, thus reducing customer waiting time virtually to zero. Similarly, an accounting firm may be able to respond quickly to customer demand by having a very large number of accountants on standby waiting for demand that may occur. An alternative perspective on speed is the time taken for goods and services to move through the chain. So, for example, products that move quickly down a supply chain from raw material suppliers through to retailers will spend little time as inventory because to achieve fast throughput time, material cannot dwell for significant periods as inventory. This in turn reduces the working capital requirements and other inventory costs in the supply chain, so reducing the overall cost of delivering to the end customer. Achieving a balance between speed as responsiveness to customers' demands and speed as fast throughput (although they are not incompatible) will depend on how the supply chain is choosing to compete.

Dependability

Dependability in a supply chain context is similar to speed in so much as one can almost guarantee 'on-time' delivery by keeping excessive resources within the chain. However, dependability of throughput time is a much more desirable aim because it reduces uncertainty within the chain. If the individual operations in a chain do not deliver as promised on time, there will be a tendency for customers to over-order, or order early, in order to provide some kind of insurance against late delivery. The same argument applies if there is uncertainty regarding the *quantity* of products or services delivered. This is why delivery dependability is often measured as 'on-time, in-full' in supply chains.

Flexibility

In a supply chain context, flexibility is usually taken to mean the chain's ability to cope with changes and disturbances. Very often this is referred to as supply chain agility. The concept of agility includes previously discussed issues such as focusing on the end customer and ensuring fast throughput and responsiveness to customer needs. But, in addition, agile supply chains are sufficiently flexible to cope with changes, either in the nature of customer demand or in the supply capabilities of operations within the chain.

Cost

In addition to the costs incurred within each operation to transform its inputs into outputs, the supply chain as a whole incurs additional costs that derive from each operation in a chain doing business with each other. These 'transaction' costs include the costs of finding appropriate suppliers, setting up supplier agreements (e.g. tendering, negotiation and contracting), running ongoing supply (e.g. ordering, expediting, monitoring supply performance, invoicing and payment), dealing with failure, supplier training, and potentially the costs of exiting an unsatisfactory relationship. It is important to consider the transaction costs of trading with a supplier as opposed to simply the price of the service or product, because they can add substantially to the *total cost* of trading. Many of the recent developments in supply chain management, such as partnership agreements or reducing the number of suppliers, are an attempt to minimise such transaction costs.

Sustainability

Any organisation that subscribes to sustainability objectives will want to make sure that it purchases its input products and services from suppliers that are similarly responsible. This may involve, for example, buying from local suppliers where possible, sourcing supplies from suppliers with ethical practices, choosing environmentally friendly products and services, using minimal packaging and transporting products by ground transport rather than air.

Efficient versus responsive supply networks

As discussed in Chapter 1, different products or services often exhibit clear differences in how they compete in markets. One popular approach used in supply chain management is to distinguish between products or services that are 'functional' and those that are 'innovative'.

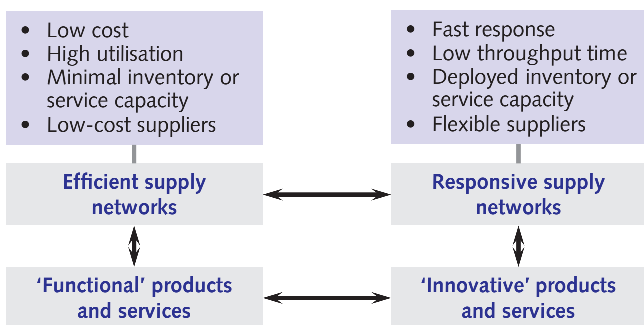
Demand for functional products and services is relatively stable and predictable, while demand for innovative products and services will be far more uncertain. In addition, the profit margins are typically much higher for innovative as opposed to functional offerings. Even within the same company, there may be *both* functional and innovative categories in evidence. For example, some of the work carried out by consultants is very standardised, with just small variations from one client to the next, while other work is highly tailored for each specific project. Hospitals have routine 'standardised' surgical procedures, such as cataract removal, but also have to provide very customised emergency post-trauma surgery. Shoe manufacturers may sell classics that change little over the years, alongside fashion shoes that last a single season.

Depending on the nature of the product or service, different supply networks are likely to be more suitable (an idea originally proposed by Professor Marshall Fisher).[2] For functional products and services, efficient (or lean) supply chain policies are typically most appropriate. This includes keeping inventories or service capacity low, especially in the downstream parts of the network, so as to maintain fast throughput and reduce the amount of working capital tied up in inventory or wasted service capacity. There is also a significant focus on maximising utilisation of all resources in the supply network to minimise costs. Information must flow quickly up and down the chain so that schedules can be given to maximise the amount of time to adjust efficiently. The chain is then managed to make sure that products or services flow as quickly as possible down the chain.

In contrast, innovative products or services are more suited to responsive (or agile) supply chain policies. Here, the emphasis is on high service levels and responsive supply to the end customer. The inventory or service capacity in the network will be deployed as close as possible to the customer. In this way, the network is still able to supply even when dramatic changes occur in customer demand. Fast throughput from the upstream parts of the network will still be needed to replenish downstream with product or service operations. Figure 7.3 outlines some of the key characteristics of functional versus innovative products and services. It also illustrates the natural alignment between these characteristics and the type of supply network that is most appropriate. In practice, some firms take a more hybrid

OPERATIONS PRINCIPLE

'Functional' products and services suit efficient (lean) supply chain management; 'innovative' products and services suit responsive (agile) supply chain management.



Example	Bucket	Bread	Mobile phone	Fashion bag
Time between new product/ service introductions	10 yr+	1 yr–10 yr+	1 yr–18 months	3–6 months
Profit margins	Tiny	Small	Very high	High
Volume and variety	High/very low	High/low	Moderate/ moderate	Moderate/ moderate
Demand volatility and uncertainty	Very low	Very low	Moderate	Moderate–high

Figure 7.3 Aligning product and service characteristics with supply network design

approach in operating their supply networks. For example, Inditex, one of the world's largest and most successful fashion retailers, operates a predominantly agile supply network, but still has aspects of lean within some of its activities. We provide more detail on lean, agile and so-called 'leagile' in Chapter 11.

7.2 Diagnostic question: *What type of supply chain relationship should be adopted*?

The 'relationship' between operations in a supply chain is the basis on which the exchange of products, services, information and money is conducted. As such, managing supply networks is about managing relationships. In considering the type of supply chain relationship, we can consider *what* the operation outsources and *who* supplies the products or services. In Chapter 4 we examined the question of *what* to outsource (from everything at one extreme to nothing at the other). In considering *who* is chosen to supply services or products, managers must decide how many suppliers to use and how close the relationship should be. Figure 7.4 illustrates this way of characterising relationships.

'Transactional' versus 'partnership' relationships

Transactional (or contractual) relationships involve purchasing products and services in a 'pure' market fashion, often seeking the 'best' supplier every time it is necessary to make a purchase. Transactional relationships can be either long or short term, but there is no guarantee of anything beyond the immediate contract. They are appropriate when short-term benefits are important. In contrast, partnership (or collaborative) relationships are longer term and involve a commitment to work together over time to gain mutual advantage. These relationships emphasise cooperation, frequent interaction, information sharing, joint problem solving and sometimes even profit sharing. The concept of mutuality is important here. A supplier does not

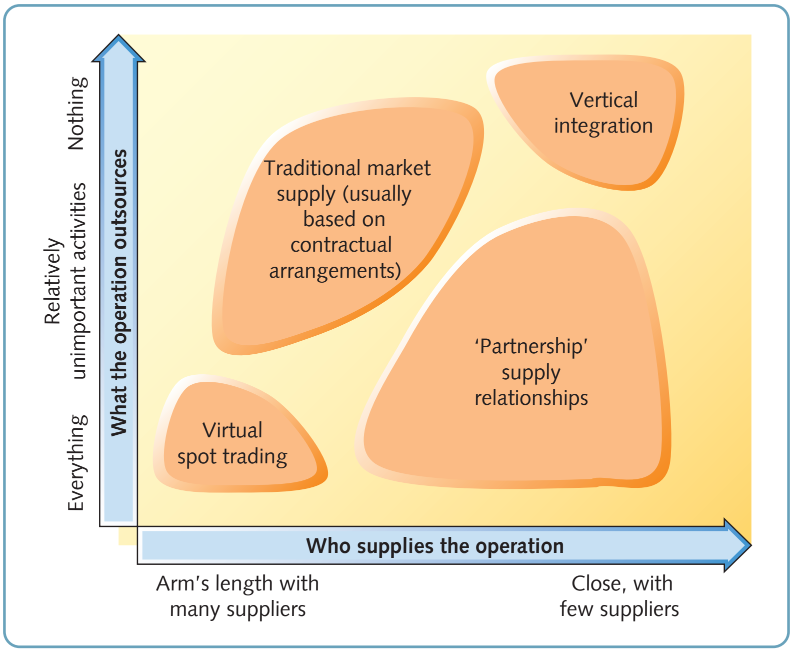


Figure 7.4 Types of supply network arrangement

become a 'partner' merely by being called one. True partnerships imply mutual benefit, and often mutual sacrifice. If it is not in the culture of a business to give up some freedom of action, it is very unlikely to ever make a success of partnerships. Table 7.1 outlines some of the key advantages of the two alternative extremes of supply chain relationships.

It is very unlikely that any business will find it sensible to engage exclusively in one type of relationship or another. Most businesses will have a portfolio of, possibly, widely differing relationships. In addition, there are degrees to which any particular relationship can be managed on a transactional or partnership basis. The real question is: where, on the spectrum from transactional to partnership, should each relationship be positioned? While there is no simple formula for choosing the 'ideal' form of relationship, there are some important factors that can sway the decision. The most obvious issue will concern how a business intends to compete in its market-place. If price is the main competitive factor, then the relationship could be determined by which approach offers the highest potential savings. On one hand, market-based contractual relationships could minimise the actual price paid for purchased services, while partnerships could minimise the transaction costs of doing business. If a business is competing primarily on product or service innovation, the type of relationship may depend on where innovation is likely to happen. If innovation depends on close collaboration between supplier and customer, partnership relationships are needed. On the other hand, if suppliers are busily competing to out-do each other in terms of their innovations, and especially if the market is turbulent and fast growing (as with many software and internet-based industries), then it may be preferable to retain the freedom to change suppliers quickly using market mechanisms.

OPERATIONS PRINCIPLE

All supply chain relationships can be described by the balance between their 'contractual' and 'partnership' elements.

Table 7.1 Benefits of transactional and partnership supply chain relationships

Transactional relationship benefits	*Partnership relationship benefits*
– Retains competition between alternative suppliers	– Higher levels of loyalty to customers than transactional relationships
– If demand changes, customer can change the number and type of suppliers; a faster and cheaper alternative to redirecting internal activities	– Reduced time and effort of frequent re-contracting
	– Reduced transaction costs
	– Reduced cost of compliance monitoring
– Wider variety of innovation sources (though access may be harder than in partnership relationships)	– Fewer quality failures and unanticipated failures
	– Earlier failure identification
– Useful in assessing a supplier before potential move to partnership model	– Emphasis on joint problem solving during failure episodes, rather than assigning blame
– Good for one-off or irregular purchases	– Generation of increased value by leveraging shared competences

Case example

Donkeys – the unsung heroes of supply networks[3]

The last 150 years has witnessed huge changes in the way that products are transported across supply networks. In shipping, 'containerisation' has enabled previously unthinkable levels of efficiency across maritime operations, paving the way for globalisation. The same period saw the Wright brothers make the first successful powered flight in 1903, ultimately leading to the creation of the aviation sector and air freight as a new mode of transportation. Similarly, Karl Benz's 'motorwagen' patent in 1886 eventually led to a new dominant mode of ground-based transportation. Yet, throughout all this turbulence, there remains an unsung hero of supply networks around the world – the humble donkey.

For thousands of years, people have used animals of all shapes and sizes to move their possessions and supplies from one location to another, including camels, dogs, elephants, goats, llamas, oxen, reindeer and sheep to name a few. But by far the most widely used animal in modern supply chains is the donkey (alongside ponies, mules or horses they are collectively known as 'packhorses'). Around the world, donkeys fulfil a wide range of roles, acting as trucks, ambulances, tractors and school buses. They carry sugar cane in Peru, transport crops in Argentina, move flowers in Jamaica and maize in Malawi, pull snow sledges in the USA, collect rubbish in Mexico, serve as a mobile beach bar in Brazil, carry fish in Mauritania, collect water in Senegal, move families in Namibia, pull mobile libraries in Ethiopia, haul building materials in Kenya, lead livestock in Yemen, are ambulances for women in labour in Afghanistan, work in India's brick kilns and plough land in China. In many tourist destinations around the world, donkeys, mules, ponies and horses are used to pull tourist carriages or to give children rides on beaches.

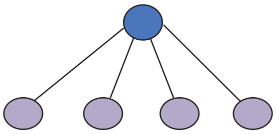
The decision to use a donkey for transportation follows the same fundamental process as any other mode of transport. First, transport options must be identified. These choices are naturally constrained by the terrain over which products need to be moved and the existing supply chain infrastructure. Second, judgements must be made concerning the relative performance benefits (in terms of quality, speed, dependability, flexibility, cost and sustainability) of available methods. As such, even when an alternative might be *technically* possible, the donkey may still represent the best balance of supply chain performance. This helps explains why, in a world of increasingly advanced technological solutions, there remains an important place for the donkey and its packhorse cousins. Yet, despite their role in helping build civilisation as we know it, packhorses are sadly some of the most mistreated animals on the planet. Cruelty comes in the form of malnutrition, beatings, overloading, poorly fitting packs and long hours in extreme climates. A number of animal charities around the world have lobbied to establish codes of conduct for animals involved in transportation. In addition, some charities provide access to cheap or free veterinary treatment, as well as offering advice to owners on how to care for their animals.

pratan ounpitipong/Moment/Getty Images

7.3 Diagnostic question: *Has the sourcing configuration been determined?*

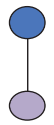
After the supply chain relationship that will be adopted for different groups of services or products has been determined, managers must consider their preference in relation to sourcing configuration. This involves selecting appropriate sourcing approaches (multiple sourcing, single sourcing, delegated sourcing or parallel sourcing) in relation to different categories of purchasing spend (non-critical, bottleneck, leverage and strategic).

Sourcing approaches



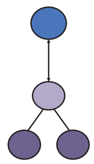
Multiple sourcing

Multiple sourcing involves obtaining a product or service component from more than one supplier. It is commonly seen in competitive markets where switching costs are low and performance objectives are primarily focused on price and dependability. Multiple sourcing can help maintain competition in the supply market, reduce supply risk and increase flexibility in the face of supplier failure or changes in customer demand. It can also allow the buyer to tap into a more diverse range of knowledge and expertise. In addition, some firms like to multi-source to prevent supplier dependence, thus allowing for changes in purchase volumes without the risk of supplier bankruptcy. However, the disadvantage of multiple sourcing is that it requires effort to ensure consistent service quality. In addition, it becomes hard to encourage supplier commitment with suppliers who are less willing to invest in new processes in the absence of longer-term contracts.



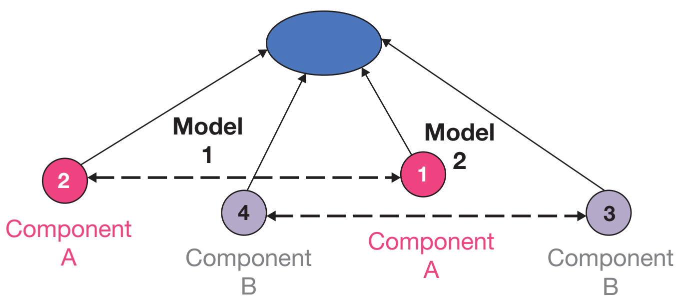
Single sourcing

Single sourcing involves buying all of one product or service component from a single supplier. Often these components represent a high proportion of total spend and are of strategic importance. In other cases, however, firms simply prefer the simplicity (and reduced transaction costs) of single sourcing. Many single-source arrangements have a longer-term outlook and focus on a wider range of performance objectives. The mutual dependency effects of single sourcing can encourage more commitment and effort, and ultimately higher-quality service or product provision. However, single-source arrangements can carry an increased risk of lock-in, upward price pressure and a reduction in the firm's bargaining power. In addition, such relationships are vulnerable to disruption if a failure to supply occurs.



Delegated sourcing

Delegated sourcing involves a tiered approach to managing supplier relationships. This means that one supplier is responsible for delivering a package of services as opposed to an individual service, or an entire sub-assembly as opposed to a single part. This has the advantage of reducing the number of first-tier suppliers significantly while simultaneously allowing a focus on strategic partners. However, delegated sourcing can alter the dynamics of the supply market and risk creating 'mega-suppliers' with significant power in the network.



Parallel sourcing

Parallel sourcing has the aim of providing the advantages of both multiple sourcing and single sourcing simultaneously. It involves having single-source relationships for a single component for different service packages or product models. If a supplier is deemed unsatisfactory, it is possible to switch to the alternative supplier who currently provides the *same* component but for a *different* service package or product model. The advantage of this sourcing approach is that it maintains competition and allows for switching in circumstances of supplier failure. However, managing delegated-sourcing arrangements is a relatively complex task.

Supply base reduction

The last 25 years has seen a trend for organisations to reduce their supplier base. This trend, building on the idea of partnership relationships, is partly an acknowledgment that organisations have limited resources and therefore some will opt to focus on developing fewer but higher-quality, relationships with key suppliers. Supply base reduction often results in significant reductions in daily running costs, such as the costs of ordering, expediting, supplier visits and various failure interventions. However, there are also potential downsides, notably around increased supply risk with some suppliers more able to act opportunistically. In addition, power dynamics typically shift in favour of the suppliers as the buyer becomes increasingly dependent on its remaining suppliers.

Making the sourcing strategy decision

> **OPERATIONS PRINCIPLE**
> ***Sourcing strategy is governed partly by the risk in the supply market, and partly by the criticality of the item to the business.***

A key challenge is to decide which sourcing strategy is most suitable under different circumstances. Here we can explore two key questions – *what is the risk in the supply market*, and *what is the criticality of the product or service to the business?* Considering risk, key issues are typically the number of alternative suppliers, how easy it is to switch from one supplier to another, exit barriers, and the cost of bringing operations back in-house. For criticality, mangers typically consider a product or service component's importance in terms of percentage of total purchase cost or the impact on business growth. By looking at these two dimensions, it is possible to position product or service components broadly in one of four key quadrants – *non-critical*,

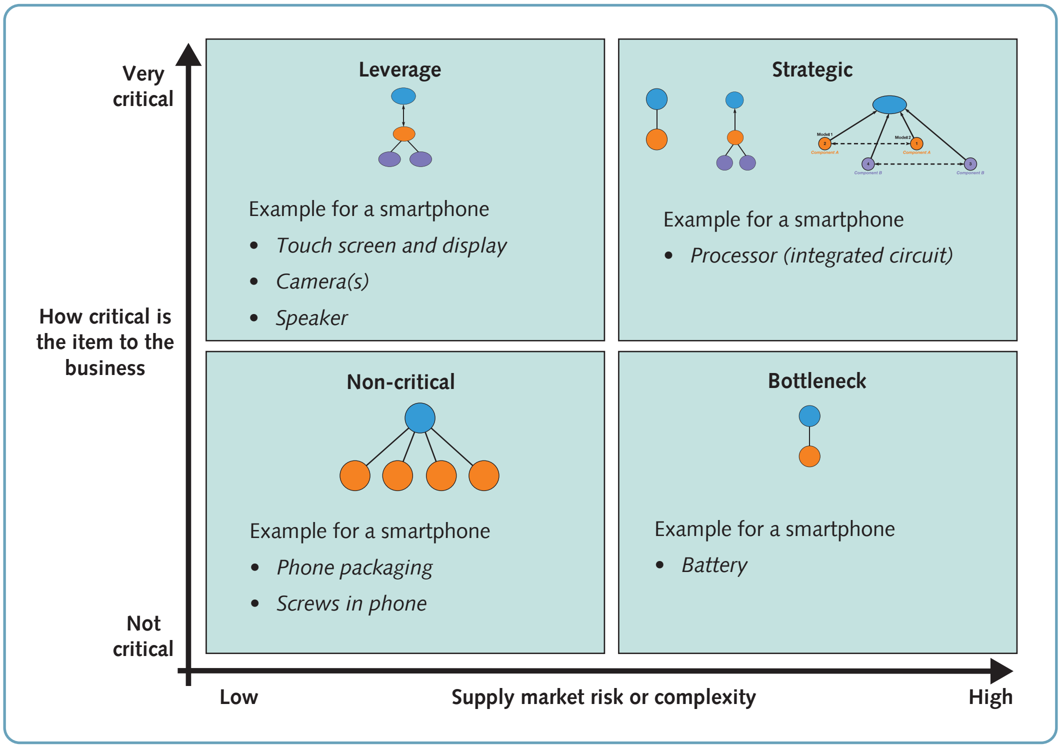


Figure 7.5 **Key sourcing groups for a smartphone manufacturer**

bottleneck, *leverage* and *strategic.*[4] Figure 7.5 provides examples of these categories for a smartphone manufacturer and outlines favoured sourcing approaches for each quadrant:

Non-critical: The packaging for transportation and display, and the screws that hold the phone components together, account for a relatively low proportion of the total cost of the product. In addition, with a large number of alternative suppliers, the supply risk is low. For the non-critical quadrant, multiple-sourcing strategies tend to be most common, though supply base reduction initiatives sometimes see single-supplier arrangements but with short contract terms.

Bottleneck: The power pack for a smartphone is relatively low cost compared to other components that make up the product. However, the limited supply alternatives and relatively high switching costs increase supply risk. For products and services in the bottleneck category, single sourcing is common because of a lack of choice in the supply market. In addition, firms sometimes look to reduce the specificity of their requirements and, in doing so, increase the number of supplier options available to them.

Leverage: The touch screen and display, and to a lesser extent camera(s) and speaker, account for a high proportion of the purchasing cost of the smartphone. However, these components are easier to source as there are a relatively large number of available (and reliable) suppliers. Suppliers in the leverage quadrant typically need to be price competitive, given the strong bargaining position that abundant supply gives to the buyer. For leverage products and services, bundling of requirements allows a shift towards delegated sourcing in many cases.

Strategic: Strategic products or services are both complex to acquire and critical to the business, accounting for a significant proportion of total spending. In this example, the processor (the 'brain' of the smartphone) sits in the strategic quadrant. There are relatively few firms capable of supplying components to sufficient quality and so the cost of switching is high. For strategic products and services, single-sourcing approaches remain popular. However, given the associated risks of single sourcing noted earlier, some firms have moved to delegated- or parallel-sourcing approaches.

7.4 Diagnostic question: *Has the approach to supplier selection and negotiation been determined?*

Supplier selection is vital for effective supply chain management. Poorly performing suppliers can reduce the performance of all the supply chains of which they are a part. Supplier selection involves initial qualification, identifying key measurement criteria, collecting information on potential suppliers, making a final selection and negotiating with suppliers to agree the terms of the engagement.

Initial qualification of suppliers

The initial qualification of suppliers is aimed at reducing possible suppliers to a manageable set prior to subsequent assessment. Pre-qualification criteria are extremely varied, but they typically focus on establishing a minimum threshold against three capabilities:

1. ***Technical capability*** – does the supplier have service or product knowledge to supply to the required specification?
2. ***Operations capability*** – does the supplier have the process knowledge to ensure consistent, responsive, dependable and reasonable-cost supply?
3. ***Financial capability*** – does the supplier have the financial strength to fund the business in both the short and long term?

Determining measurement criteria

Not all performance objectives will have the same importance, so when selecting suppliers, the relative importance of key performance objectives (quality, speed, dependability, flexibility, cost, sustainability and others) must be determined. After this, measurable criteria for key performance objectives need to be determined. For example, for cost a firm might consider unit price, pricing terms (e.g. volume discounts), exchange rate effects, and so on. Similarly, for delivery it might consider average lead-time, delivery frequency, minimum order size, distance from its premises and percentage of on-time deliveries based on verifiable data.

OPERATIONS PRINCIPLE
Measurement criteria for supplier selection should align with the overall performance objectives of the supply network.

Obtaining relevant information on prospective suppliers

As the group of potential suppliers narrows, additional levels of detail may be collected via requests for quotes (RFQs) or requests for information (RFIs). In addition, visits to the suppliers and tests to assess supplier competence, such as trial orders, may be conducted. The amount of time and effort invested in information search is partly influenced by the strategic importance of the purchase and the perceived capability of the supply base. For example, when the supply market is generally considered to be capable and the strategic importance of the purchase is low, limited information search is appropriate. Conversely, if the firm is buying a service or product of strategic importance and the supply market is more uncertain, more investment in information search will be necessary. The type of purchase also influences the information search process. Low levels of information search are needed for *repeat or routine re-buys* (for example, placing an order for a well-established service or product with an existing supplier). *Modified re-buys*, (for example, where new services are bought from known suppliers, or where existing services are bought from new suppliers) have moderate uncertainty and hence justify moderate information search. Finally, *new buys* (for example, sourcing entirely new services or products from unknown suppliers) have high levels of uncertainty and require an extensive information search.

Making the selection

Having arrived at a group of viable alternatives, rarely is one potential supplier so clearly superior that the decision is self-evident. Therefore, alternative attributes need trading off, usually by using some kind of supplier 'scoring' or assessment procedure to rate alternative suppliers. These multi-criteria decision-making models include the weighted score method (see worked example below) and analytical hierarchy process (AHP). Such models aim to provide quantifiable information for each key selection criterion and a weighting of their relative importance to allow for an objective assessment of different suppliers. Total cost of ownership (TCO) is an alternative approach that seeks to provide detailed information on all possible costs (rather than simply the price of the service or product) associated with procurement to reach more 'rational' decisions during supplier selection.

Negotiating with suppliers during supplier selection

A key part of the selection process is negotiation. The approach to negotiation is naturally affected by initial decisions around performance priorities, supplier type (i.e. traditional versus partnership) and sourcing configuration (e.g. single, multi, delegated or parallel). However, regardless of the general approach, managers who are well informed around

Worked example

Weighted score method in supplier selection

A commercial bank in Belgium has decided to change its legal services provider as it has become dissatisfied with its current supplier. The three legal firms have been evaluated using a 1–10 scale against a set of key criteria. Each of these criteria has a weighting, again from 1–10, signifying its relative importance to the bank in its purchasing decision. Table 7.2 illustrates that, based on this assessment, Juris Civilis is the preferred supplier. Assuming that subsequent negotiations reach a satisfactory conclusion, they would be selected by the bank to provide legal services.

Table 7.2 Weighted supplier selection criteria for legal services provider

Selection criteria	*Weight*	*Sullivan & Anderson*	*Juris Civilis*	*Altium Legal*
Quality of service	10	7 (10 × 7 = 70)	8 (10 × 8 = 80)	6 (10 × 6 = 60)
Range flexibility	8	5 (8 × 5 = 40)	7 (8 × 7 = 56)	6 (8 × 6 = 48)
Volume flexibility	4	8 (4 × 8 = 32)	3 (4 × 3 = 12)	7 (4 × 7 = 28)
Average service speed	7	5 (7 × 5 = 35)	8 (7 × 8 = 56)	9 (7 × 9 = 63)
Record of dependability	9	4 (9 × 4 = 36)	8 (9 × 8 = 72)	7 (9 × 7 = 63)
Sustainability	4	5 (4 × 5 = 20)	6 (4 × 6 = 24)	7 (4 × 7 = 28)
Potential innovation	3	5 (3 × 5 = 15)	7 (3 × 7 = 21)	7 (3 × 7 = 21)
Price of service	6	9 (6 × 9 = 54)	6 (6 × 6 = 36)	5 (6 × 5 = 30)
Total weighted score		**302**	**357**	**341**

Table 7.3 Common tactics used in buyer–supplier negotiations

Tactic	*Considerations for using negotiation tactic*
EMOTION	• Positive and negative emotions are extremely influential in 'framing' the negotiation • Anger is the most common emotion in negotiation, but also the most damaging to potential settlement • Positive emotions support collaborative approaches to negotiation • Use emotions early in negotiation to maximise influence • Be sincere and don't let emotions control you • Use to increase 'perceived value' when bargaining
LOGIC	• Ensure your logic is credible and use one powerful argument rather than multiple weaker arguments • Try to get your logic in before the other negotiating party • Use logic to counter emotion tactics
THREAT	• Avoid threatening where possible and be credible with any threat • If you do use threat, threaten the business rather than the person • Discreet or veiled threats are less risky than explicit threats • Use 'if' in any threat to give the opponent room to manoeuvre
BARGAINING	• Avoid exposing your position too early in negotiations • Move in small and decreasing steps and extract return for any movement where possible • Don't move too quickly and avoid appearing too keen to move
COMPROMISE	• Use compromise after other tactics • Use extreme (but credible) position to maximise value when using compromise • Avoid suggesting compromise too early in a negotiation as it shuts down more creative solutions and favours the other negotiating party

negotiation techniques are likely to see better outcomes. During negotiations between buyers and suppliers, a number of tactics are often in evidence, including *emotion*, *logic*, *threat*, *bargaining* and *compromise*. Table 7.3 outlines some important considerations when using each tactic.

While it is essential for negotiators to be skilled at the tactics described in Table 7.3, they can backfire. This 'advocacy approach', whereby a negotiator advocates for one party to gain its most favourable outcome, carries a number of risks. These include an emphasis on short-term solutions at the expense of potential longer-term gains, personalised conflicts, damage to the relationship between the buyer and supplier, and potentially a 'lose–lose' outcome. An alternative approach is 'collaborative negotiation', also called principled negotiation or mutual gains bargaining. This requires negotiators to adopt a mindset that emphasises meeting the needs of both parties, and involves a transparent process, investment in developing personal relationships, using creative brainstorming techniques (growing the pie rather than dividing a 'fixed pie'), and favouring longer-term win–win agreements.

7.5 Diagnostic question: *Are supply chains run effectively?*

So far, we have focused on *establishing* supply chain relationships. We now consider how such relationships can be managed to maximise the ongoing performance of the supply network. This involves examining information flows, perception gaps, the development of suppliers and the use of logistics (distribution).

Managing information flows

Managing supply chain relationships is not just a matter of choosing the right suppliers and then leaving them to get on with day-to-day supply. It is also about ensuring that suppliers are given the right information and encouragement to maintain smooth supply. Some mechanism should ensure a two-way flow of information between customer and supplier. Both suppliers and customers can forget to communicate developments that could affect the other. Customers may see suppliers as having the responsibility for ensuring supply 'under any circumstances'. Conversely, suppliers may be reluctant to inform customers of potential problems with supply because it may risk the relationship. Yet, if customer and supplier see themselves as 'partners', the free flow of information and a mutually supportive tolerance of occasional problems is the foundation of smooth supply. Often day-to-day supplier relationships are damaged because of internal inconsistencies. For example, one part of a business may be asking a supplier for some service beyond the strict nature of their agreement, while another part of the business is late in paying the same supplier.

Service-level agreements

Some organisations bring a degree of formality to supplier relationships by encouraging (or requiring) all suppliers to agree service-level agreements (SLAs). SLAs are formal definitions of the dimensions of service and the relationship between suppliers and the organisation. The types of issues covered by such an agreement could include response times, the range of services, dependability of service supply, and so on. Boundaries of responsibility and appropriate performance measures could also be agreed. Taking an internal supply network example, an SLA between an information system support unit and a research unit in the laboratories of a large pharmaceutical company could define such performance measures as:

- the types of information network services that may be provided as 'standard';
- the range of special information services that may be available at different periods of the day;
- minimum 'up time', i.e. the proportion of time the system will be available at different periods of the day;
- maximum response time and average response time to get the system fully operational should it fail;
- maximum response time to provide 'special' services.

Although SLAs are described here as mechanisms for governing the ongoing relationship between suppliers and customers, they often prove inadequate because they are seen as being useful in setting up the terms of the relationship, but then are used only to resolve disputes. For SLAs to work effectively, they must be treated as working documents that establish the details of ongoing relationships *in the light of experience*. Used properly, they are a repository of the knowledge that both sides have gathered through working together. Any SLA that stays unchanged over time is failing to encourage improvement in supply.

Perception differences in supply chain relationships

One of the biggest issues for successful SCM is the mismatch between how customers and suppliers perceive both what is required and how the relationship is performing. Figure 7.6 illustrates the four key gaps that sometimes emerge in supply chain relationships. Taking the *customer perspective* (Scenario A in Figure 7.6) first, you (the customer) have an idea about

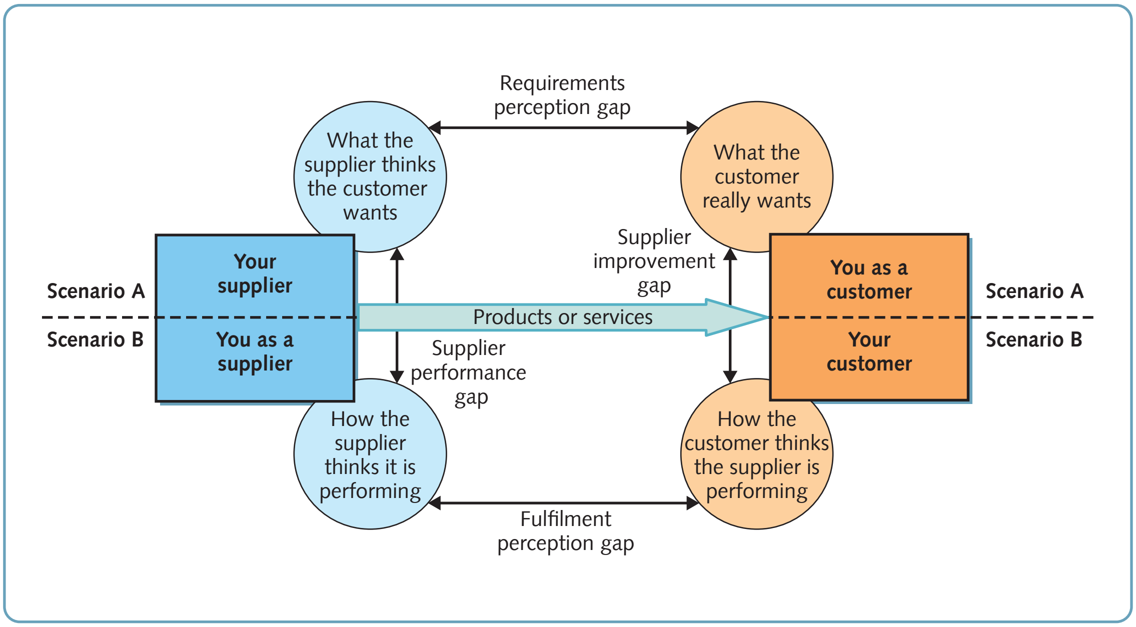


Figure 7.6 **Potential perception mismatches in supply chain relationships**

what you really want from a supplier. This may, or may not, be formalised in the form of a service level agreement. But no SLA can capture everything about what is required. There may be a gap between how you as a customer describe what is required and how the supplier interprets it. This is the *requirements perception gap*. Similarly, as a customer, you will have a view on how your supplier is performing in terms of fulfilling your requirements. Misalignment between your view and how your supplier believes it is performing is evidence of a *fulfilment perception gap*. Both these gaps are a function of the effectiveness of the communication between supplier and customer. Third, the *supplier improvement gap* is the gap between what the customer wants from its supplier and how it perceives the supplier's performance. This will influence the kind of supplier development goals set by the customer. Finally, the *supplier performance gap* is the gap between your supplier's perceptions of your needs and its own assessment of performance. This will indicate how a supplier initially sees themselves improving their own performance.

From a *supplier perspective* (Scenario B in Figure 7.6), the same approach can be used to understand *customer* perceptions, both of their requirements and their view of your (the supplier's) performance. What is less common, but can be equally valuable, is to use this model to examine the question of whether customer requirements and perceptions of performance are either accurate or reasonable. For example, customers may be placing demands on suppliers without fully considering their consequences. It may be that slight modifications in what is demanded would not inconvenience the customers and yet would provide significant benefits to the supplier that could then be passed on to the customer. Similarly, a customer may be incompetent at measuring supplier performance, in which case current good service may not be recognised.

OPERATIONS PRINCIPLE
Unsatisfactory supply chain relationships can be caused by requirement, fulfilment, supplier improvement and supplier performance perception gaps.

Supply chain relationships are multi-tiered configurations

Up to this point, we have largely treated supply chain relationships as if they simply exist between two entire organisations. However, supply chain relationships are multi-tiered configurations, as illustrated in Figure 7.7. As such, a formal contract between two organisations

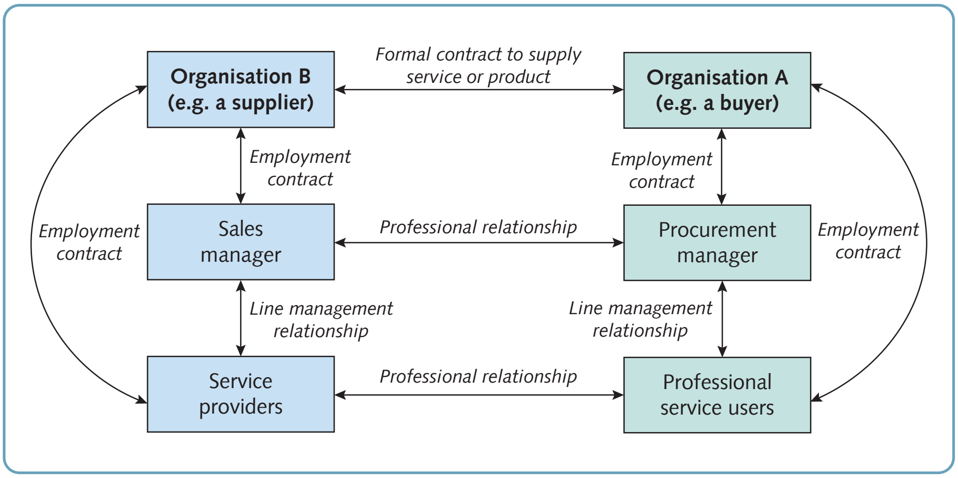


Figure 7.7 Supply chain relationships are multi-tiered configurations

Source: Developed in collaboration with Juri Matinheikki and Katri Kauppi of Aalto University, Finland, and Erik van Raaij of Erasmus University Rotterdam, Netherlands

in a supply network has to be interpreted by managers in these organisations. In turn, the service itself is then delivered by service personnel and received by service users. At each of these levels, there is essentially a relationship. Ideally, these relationships will be consistent in terms of attitudes, actions and perceptions. But in reality, there are often significant differences across these levels of supply chain relationship. Furthermore, we also see complex relationship dynamics between the organisation and its employees. Again, disconnects between what the organisation wants its managers and employees to deliver and how the service is actually delivered on the ground can create significant problems in multi-tiered supply chain relationships.

Developing suppliers

Unless a relationship is purely market-based and transactional, it is in a customer's long-term interests to take some responsibility for developing supplier capabilities. Helping a supplier to improve not only enhances the service from the supplier; it may also lead to greater supplier loyalty and long-term commitment. This is why some particularly successful businesses invest in supplier development teams whose responsibility is to help suppliers improve their own operations processes. Of course, committing the resources to help suppliers is only worthwhile if it improves the effectiveness of the supply network as a whole. The process of supplier development can be broken down into four key stages:

1. ***Select product or service and supplier for development*** – not all suppliers can be developed for all products and services, but those that provide strategic products or services (in the 'strategic' quadrant of Figure 7.5) are likely candidates. Of these, those with relatively weak current performance but with potential for improvement are likely to be appealing, as are suppliers whose inputs are particularly important, or where switching to another supplier is difficult.
2. ***Form a project team and gain buy-in for supplier development*** – the project team should bring together key stakeholders from the buyer, supplier and other relevant parties. We discuss this issue in detail within the project management chapter (see Chapter 15), but it is worth noting the importance of cooperation from key stakeholders. Gaining 'buy-in' can require significant time, sensitivity and effort in the early stages of the supplier development process, especially if the supplier's management could be relatively defensive of development initiatives.

3. ***Agree goals and measures for supplier development*** **–** like all projects, supplier development requires clear scoping in terms of timelines, costs and key deliverables to ensure that both buyer and supplier agree on what constitutes success. This stage should also involve risk evaluation so that project leaders can consider possible mitigation strategies. For example, if a lack of supplier commitment is identified as a risk, mitigation might include more investment in the internal marketing of potential benefits or refined financial incentives (such as profit sharing).

4. ***Implement, monitor and learn*** **–** it is important to monitor progress and intervene if the project is deviating from its performance targets. Similarly, it is important to celebrate success as this helps to reinforce the value of supplier development to all stakeholders. But supplier development does not stop when a project is complete. Instead, the objective is to develop suppliers' capabilities to enable them to continue improving without the active intervention of the customer. In addition, learning from supplier development initiatives can be extremely powerful in informing subsequent projects with other suppliers.

Logistics

Logistics (or distribution) is the activity of moving products from suppliers to their customers. For many firms, logistics is not something they consider, because they either focus solely on non-physical outputs or the products they do deliver are such a small proportion of their business that they simply use distribution services on an ad hoc basis. However, for other more product-oriented firms, managing logistics can be a critical issue. This is especially the case when the costs of distribution account for a large proportion of their total costs.

Some organisations operate *first-party logistics* (1PL), whereby the logistics activity is an entirely internal process. For example, a manufacturing firm will deliver directly, or a retailer such as a supermarket will collect products from a supplier. Some firms decide to outsource logistics services over a specific segment of a supply chain. This is known as *second-party logistics* (2PL). For example, they may hire a maritime shipping company to transport, and if necessary store, products from a specific collection point to a specific destination. *Third-party logistics* (3PL) is when a firm contracts a logistics company to work with other transport companies to manage their logistics operations more fully. It is a broader concept than 2PL and can involve transportation, warehousing, inventory management and even packaging or re-packaging products. *Fourth-party logistics* (4PL) is a yet broader idea. Accenture, the consulting group, originally termed 4PL as 'an integrator that assembles the resources, capabilities, and technology of its own organisation and other organisations to design, build and run comprehensive supply chain solutions'. 4PL firms can manage all aspects of a client's supply network and may act as a single interface between the client and multiple logistics service providers. Finally (almost inevitably), some firms are selling themselves as *fifth-party logistics providers* (5PL), mainly by defining themselves as broadening their scope further to include e-business.

Volume, size and value

In selecting methods of transportation for logistics, organisations typically consider the volume, size and value characteristics of their products. For example, a firm distributing small volumes of small and high-value products around the world is much more likely to adopt air freight, whereas a firm distributing high volumes of bulky and low-value products across its supply network is more likely to use maritime transportation. Typically, the cost of moving a given product from one location to another is highest for air, then road, rail, water, and finally pipeline. In selecting the methods of transportation that are most suitable, costs of different options clearly count. However, the choice may also be influenced by 'vertical trade-offs' – benefits that may change over time as a result of improvements in one method of transportation versus another (for example, the development of a new high-speed rail link or improved fuel economy in shipping). In addition, 'lateral trade-offs' consider the balance between costs of a particular transportation method and possible benefits elsewhere. For example, Zara predominantly adopts (high-cost) air freight for its worldwide logistics, but gains significantly from faster delivery times and the ability to hold much lower inventories across its supply network.

Case example

Twice around the world for Wimbledon's tennis balls[5]

(Our thanks to Professor Mark Johnson, of Warwick Business School for this example.)

The Wimbledon 'Grand Slam' tennis tournament is the oldest and arguably the most prestigious tennis tournament in the world. Slazenger, the sports equipment manufacturer, has been the official ball supplier for Wimbledon since 1902. Yet those balls used at Wimbledon, and the materials from which they are made, will have travelled 81,385 kilometres between 11 countries and across four continents before they reach Centre Court. Professor Mark Johnson, of Warwick Business School said: *'It is one of the longest journeys I have seen for a product. On the face of it, travelling more than 80,000 kilometres to make a tennis ball does seem fairly ludicrous, but it just shows the global nature of production, and in the end, this will be the most cost-effective way of making tennis balls. Slazenger are locating production near the primary source of their materials in Bataan in the Philippines, where labour is also relatively low cost.'*

The complex supply chain is illustrated in Figure 7.8. It sees clay shipped from South Carolina in the USA, silica from Greece, magnesium carbonate from Japan, zinc oxide from Thailand, Sulphur from South Korea and rubber from Malaysia to Bataan where the rubber is vulcanised – a chemical process for making the rubber more durable. Wool is then shipped from New Zealand to the UK, where it is pressed into felt and then flown back to Bataan in the Philippines. Meanwhile, petroleum naphthalene from Zibo in China and glue from Basilan in the Philippines are brought to Bataan where Slazenger manufacture the ball. Finally, the tins, which contain the balls, are shipped in from Indonesia and once the balls have been packaged, they are sent to Wimbledon. *'Slazenger shut down the factory in the UK years ago and moved the equipment to Bataan in the Philippines'*, says Professor Johnson. *'They still get the felt from Stroud [UK], as it requires a bit more technical expertise. Shipping wool from New Zealand to Stroud and then sending the felt back to the Philippines adds a lot of miles, but they obviously want to use the best wool for the Wimbledon balls.'*

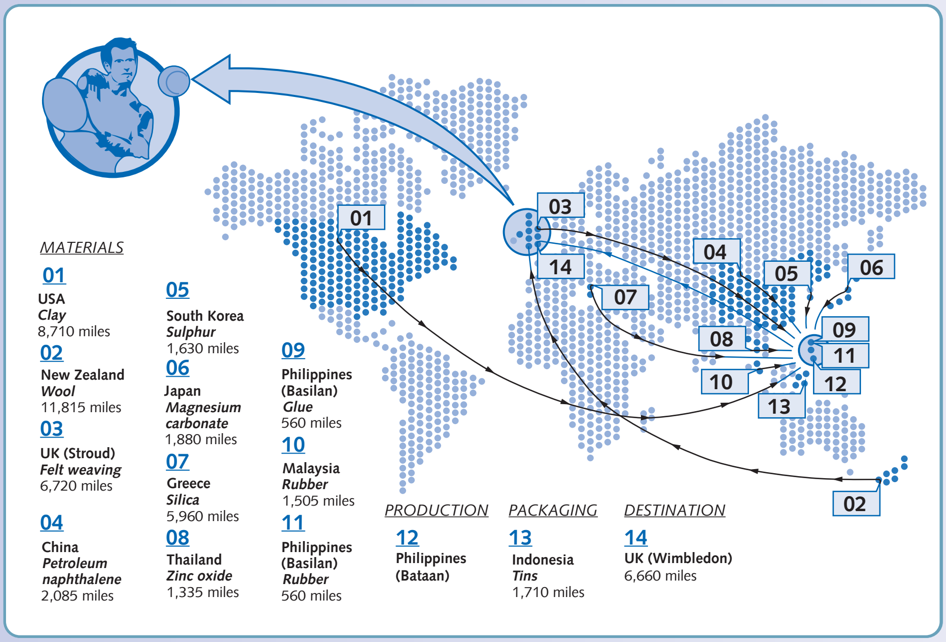


Figure 7.8 Wimbledon's tennis balls travel over 80,000 kilometres in their supply network

Source: The 50,000 mile journey of Wimbledon's tennis balls, WBS News 02/07/2014, ©2017 Warwick Business School.

Logistics and the Internet of Things (IoT)

Internet-based communication has had a significant impact on logistics and distribution, in particular through the adoption of the 'Internet of Things' (IoT). At its simplest, an IoT is a network of physical objects (such as products, equipment, materials-handling devices, trucks) that have electronics, software and sensors implanted in them that can gather and exchange data. Combined with global positioning systems (GPS) it permits instantaneous tracking of trucks, materials and people, allowing logistics companies, warehouses, suppliers and customers to share knowledge of where products are in the network and where they are going next. This permits more effective coordination and creates the potential for cost savings. For example, an important issue for transportation companies is 'back-loading'. When the company is contracted to transport goods from A to B, its vehicles may have to return from B to A empty. Back-loading means finding a potential customer who wants their goods transported from B to A in the right time frame. IoT increases the likelihood that a company can fill their vehicles on both the outward and return journeys. Similarly, IoT enables 'track-and-trace' technologies so package distribution companies can inform and reassure customers that their service is being delivered as promised.

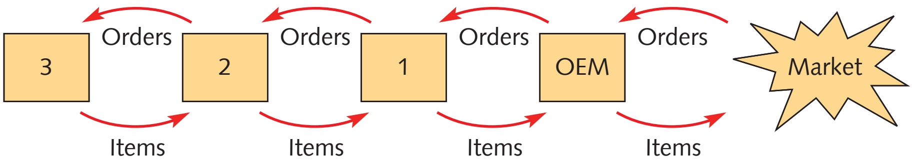
7.6 Diagnostic question: *Are supply chain dynamics under control?*

There are dynamics that exist between firms in supply chains that cause errors, inaccuracies and volatility, and these increase for operations further upstream in the supply chain. This is known as the *bullwhip effect*, so called because a small disturbance at one end of the chain causes increasingly large disturbances as it works its way 'upstream'. Its main cause is a perfectly understandable and rational desire by the different links in the supply chain to manage their levels of activity or inventory sensibly. To demonstrate this, examine the production rate and stock levels for the supply chain shown in Table 7.4. This is a simple four-stage supply chain where the focal firm is served by three tiers of suppliers. The demand from the focal firm's market has been running at a rate of 100 items per period, but in period 2, demand reduces to 95 items per period. All stages in the supply chain work on the principle that they will keep in stock one period's demand. This is a simplification but not a gross one. Many operations set their output or inventory levels to their demand rate. The column headed 'stock' for each level of supply shows the starting stock at the beginning of the period and the finish stock at the end of the period. At the beginning of period 2, the focal firm has 100 units in stock (that being the rate of demand up to period 2). Demand in period 2 is 95 and so it knows that it would need to produce sufficient items to finish up at the end of the period with 95 in stock (this being the new demand rate). To do this, it need only manufacture 90 items; these, together with five items taken out of the starting stock, will supply demand and leave a finished stock of 95 items. The beginning of period 3 finds the focal firm with 95 items in stock. Demand is also 95 items and therefore its production rate to maintain a stock level of 95 will be 95 items per period. The focal firm now operates at a steady rate of producing 95 items per period. Note, however, that a change in demand of only five items has produced a fluctuation of 10 items in the focal firm's production rate.

Carrying this same logic through to the first-tier supplier, at the beginning of period 2, the second-tier supplier has 100 items in stock. The demand that it has to supply in period 2 is derived from the production rate of the focal firm. This has dropped down to 90 in period 2. The first-tier supplier therefore has to produce sufficient to supply the demand of 90 items (or the equivalent) and leave one month's demand (now 90 items) as its finished stock.

Table 7.4 Fluctuations of production levels along supply chain in response to small change in end-customer demand

Period	*Third-tier supplier*		*Second-tier supplier*		*First-tier supplier*		*Focal firm*		*Demand*
	Production	*Stock*	*Production*	*Stock*	*Production*	*Stock*	*Production*	*Stock*	
1	100	100	100	100	100	100	100	100	100
		100		100		100		100	
2	20	100	60	100	80	100	90	100	95
		60		80		90		95	
3	180	60	120	80	100	90	95	95	95
		120		100		95		95	
4	60	120	90	100	95	95	95	95	95
		90		95		95		95	
5	100	90	95	95	95	95	95	95	95
		95		95		95		95	
6	95	95	95	95	95	95	95	95	95
		95		95		95		95	



Note: Starting stock (a) + production (b) = finishing stock (c) + demand, that is production in previous tier down (d): see explanation in text. Note all stages in the supply chain keep one period's inventory, c = d

A production rate of 80 items per month will achieve this. It will therefore start period 3 with an opening stock of 90 items, but the demand from the focal firm has now risen to 95 items. It therefore has to produce sufficient to fulfil this demand of 95 items and leave 95 items in stock. To do this, it must produce 100 items in period 3. After period 3 the first-tier supplier then resumes a steady state, producing 95 items per month. Note again, however, that the fluctuation has been even greater than that in the focal firm's production rate, decreasing to 80 items a period, increasing to 100 items a period, and then achieving a steady rate of 95 items a period. Extending the logic back to the third-tier supplier, it is clear that the further back up the supply chain an operation is placed, the more drastic are the fluctuations.

OPERATIONS PRINCIPLE
Demand fluctuations become progressively amplified as their effects work back up the supply chain.

This simple demonstration ignores a number of other factors that make the fluctuations even more pronounced. These include lack of forecasting, errors in forecasting, quantity discounting (which encourages less-frequent but larger orders), price fluctuations, time lags between the flows of information (order) and flow of materials (deliveries), variable delivery times, and panic ordering in anticipation of shortages or in reaction to them. Figure 7.9 shows the net result of all these effects in supply chains. As we move further away from the end customer, the amplitude and variance in order patterns increases significantly.

Supply chain dynamics have a number of harmful impacts on those operating in a supply network. These include the costs of outsized facilities and excess inventories to deal with demand spikes that are then often underutilised. For human resources, service capacity typically oscillates between underutilisation and overutilisation, and many firms are forced to hire,

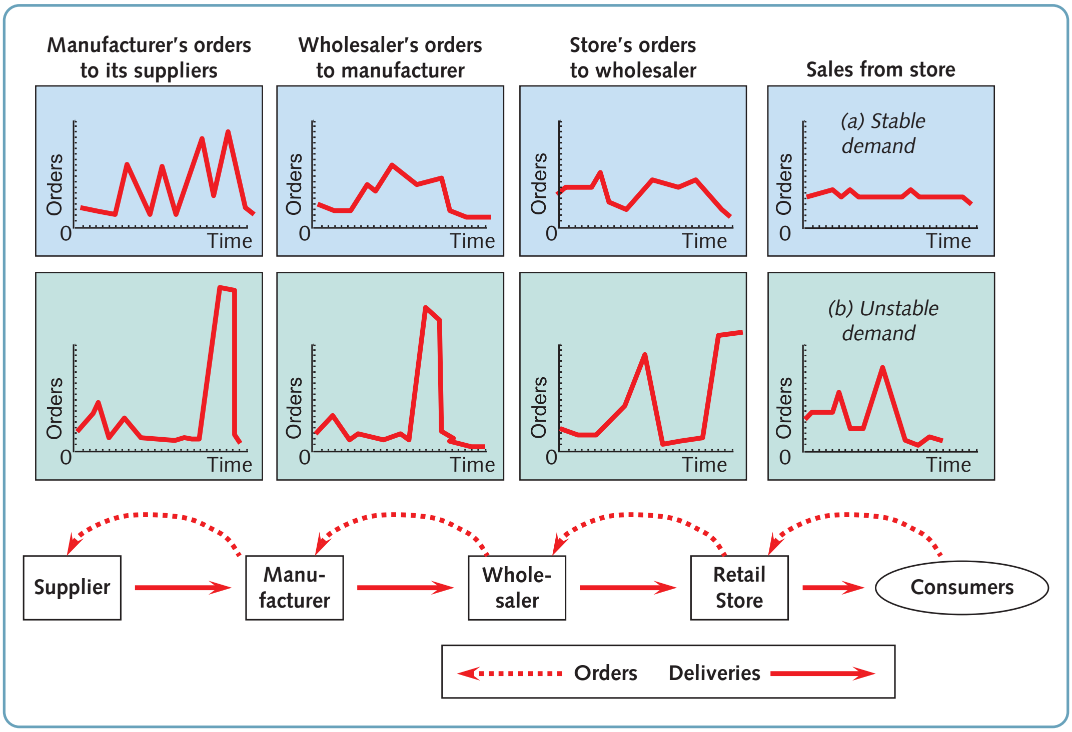


Figure 7.9 Typical supply chain dynamics for (a) stable end-customer demand and (b) unstable end-customer demand

fire and re-hire employees as they experience the volatile demand patterns caused by the bullwhip effect. In addition, irregular patterns of work cause inefficiencies, delays (with the extra costs of expediting orders) and customer dissatisfaction. Of course, this also assumes that the end-customer demand is fundamentally *stable* (see Figure 7.9(a)). Sales of services or products can be *unstable*, either because of the fundamental nature of demand, promotion activities or panic buying (as seen during the COVID-19 pandemic). Then, the bullwhip effect is even more extreme (see Figure 7.9(b)).

Controlling supply chain dynamics

An effective way to improve the performance of supply networks is to find ways to minimise the bullwhip effect. This usually means coordinating the activities of the operations in the chain in several ways.[6]

Align all the channels of information and supply

Channel alignment means the adjustment of scheduling, material movements, stock levels, pricing and other sales strategies so as to bring all the operations in the chain into line with each other. This goes beyond the provision of information. It means that the systems and methods of planning and control decision-making are harmonised through the chain. For example, even when using the same information, differences in forecasting methods or purchasing practices can lead to fluctuations in orders between operations in the chain. One way of avoiding this is to allow an upstream supplier to manage the inventories of its downstream customer. This is known as vendor-managed inventory (VMI). So, for example, a packaging supplier could take responsibility for the stocks of packaging materials held by a food manufacturing customer. In turn, the food manufacturer takes responsibility for the stocks of its products that are held in its customer's warehouses.

Increase operational efficiency throughout the supply chain

'Operational efficiency' in this context means the efforts that each operation in the chain makes to reduce its own complexity, the cost of doing business with other operations in the chain, and the supply chain's throughput time. The cumulative effect of this is to simplify the whole chain. For example, imagine a chain of operations whose performance level is relatively poor: quality defects are frequent, the lead time to order products and services is long, delivery is unreliable, and so on. The behaviour of the chain would be a continual sequence of errors and effort wasted in replanning to compensate for the errors. Poor quality would mean extra and unplanned orders being placed, and unreliable delivery and slow delivery lead times would mean high safety stocks. By contrast, a chain whose operations have high levels of operations performance is more predictable and has faster throughput, both of which help to minimise supply chain fluctuations.

Improve forecasts

Improving forecast accuracy directly reduces the inventory-holding requirements that will achieve customer service-level targets. In a service context, the more accurate the estimates of demand are, the less excess capacity is required to buffer uncertainty. In addition, reducing lead times means that you need to forecast less far into the future and thus have a large impact on bullwhip costs. Further detail on methods of demand forecasting are provided in the supplement to Chapter 4.

Share information throughout the supply chain

One of the reasons for the fluctuations in output described in the example earlier was that each operation in the chain reacted to the orders placed by its immediate customer. None of the operations had an overview of what was happening throughout the chain. If information had been available and shared throughout the chain, it is unlikely that such wild fluctuations would have occurred. It is sensible therefore to try to transmit information throughout the chain so that all the operations can monitor true demand, free of these distortions. An obvious improvement is to make information on end-customer demand available to upstream operations. Electronic point-of-sale (EPOS) systems used by many retailers attempt to do this. Sales data from checkouts or cash registers are consolidated and transmitted to the warehouses, transportation companies and supplier operations that form its supply chain. Similarly, electronic data interchange (EDI) helps to share information across supply networks and this reduces bullwhip effects. The other rapidly developing approach to gaining a trusted overview of supply chains is the use of blockchain technology (see the following section).

OPERATIONS PRINCIPLE
Supply chain dynamics (the bullwhip effect) can be reduced by aligning planning and control decisions, improving flow efficiency, better forecasting and information sharing.

Blockchain technology in supply networks

Many new technologies find particular applications for which they seem to have been designed – robots in manufacturing processing, face recognition in retail and security operations, and so on. Blockchain (technically, 'distributed ledger') technology is slightly different, having evolved from its first, and best-known, application as the accounting method for the virtual currency Bitcoin to be used in a variety of business domains, including supply chain management.

A blockchain is a decentralised, digitised, public ledger (list) of transactions (movements, authorisations, payments, etc.). A 'block' is a record of new transactions. When each block is completed (verified) it is added to the chain, thus creating a chain of blocks, or 'blockchain'. There are five basic principles underlying the technology:[7]

1. ***It uses distributed databases*** – all participants on a blockchain have access to the entire database and its complete history. No single participant controls the data or the information. Every participant can verify the records of its transaction partners directly, without an intermediary.
2. ***It uses peer-to-peer (P2P) transmission*** – all communication occurs directly between peers (or rather their computers, known as nodes) rather than through a central node. Each node in the network stores and forwards information to all other nodes.
3. ***It is transparent and can be used anonymously*** – every transaction is visible to anyone with access to the system. Each node (user) on a blockchain has a unique 30-plus-character alphanumeric address that identifies it. Users can either choose to remain anonymous, or alternatively provide proof of their identity to others. Transactions occur between blockchain addresses.
4. ***Its records are irreversible*** – once a transaction is entered in the database (and the accounts updated) the records cannot be altered, because they're linked to every transaction record that came before them (which is why it is called a 'chain'). Computational algorithms and approaches are used to ensure that the recording on the database is permanent, chronologically ordered and available to all others on the network.
5. ***It uses computational logic*** – because the distributed ledgers are digital, all blockchain transactions can be tied to a known computational logic. This means that participants can set up algorithms and rules that automatically trigger transactions between nodes (for example, to make automated payments upon delivery of services).

The role of blockchain technology in supply chain management

Blockchain technology has significant potential for supply chain management because it addresses many of the practical issues facing supply chain managers. First, supply networks are often complex and involve many operations. However, for a distributed ledger network, the larger the number of nodes, the more secure it is, so the innate complexity of supply networks actually strengthens the technology. Second, there are a very large number of transactions necessary for the smooth running of most supply networks. As such, blockchain solutions are especially useful if the supply network crosses national boundaries; customs, certification and other documents are required by many regulatory authorities (see the case example 'TradeLens – blockchain revolutionises shipping'). Third, blockchain technology is secure, with its cryptographic techniques making it practically impossible for hackers to alter information. The validation rules of blockchains mean that a potential hacker would need to access more than half of the nodes in the blockchain (which is why it is more secure to have more nodes in the blockchain). Fourth, the security of blockchains means that all parties can trust the provenance (history) of supplies. This is particularly important when dealing with food supplies, luxury goods (that are frequently forged) or ethical goods (avoiding 'blood diamonds', for example). Finally, supply chains are 'threads of operations through supply networks' – the key word being 'networks'. Blockchain technology is 'distributed' – the whole concept is based on how supply chains operate as parts of networks. Figure 7.10 illustrates how blockchains can be used in a supply chain context.

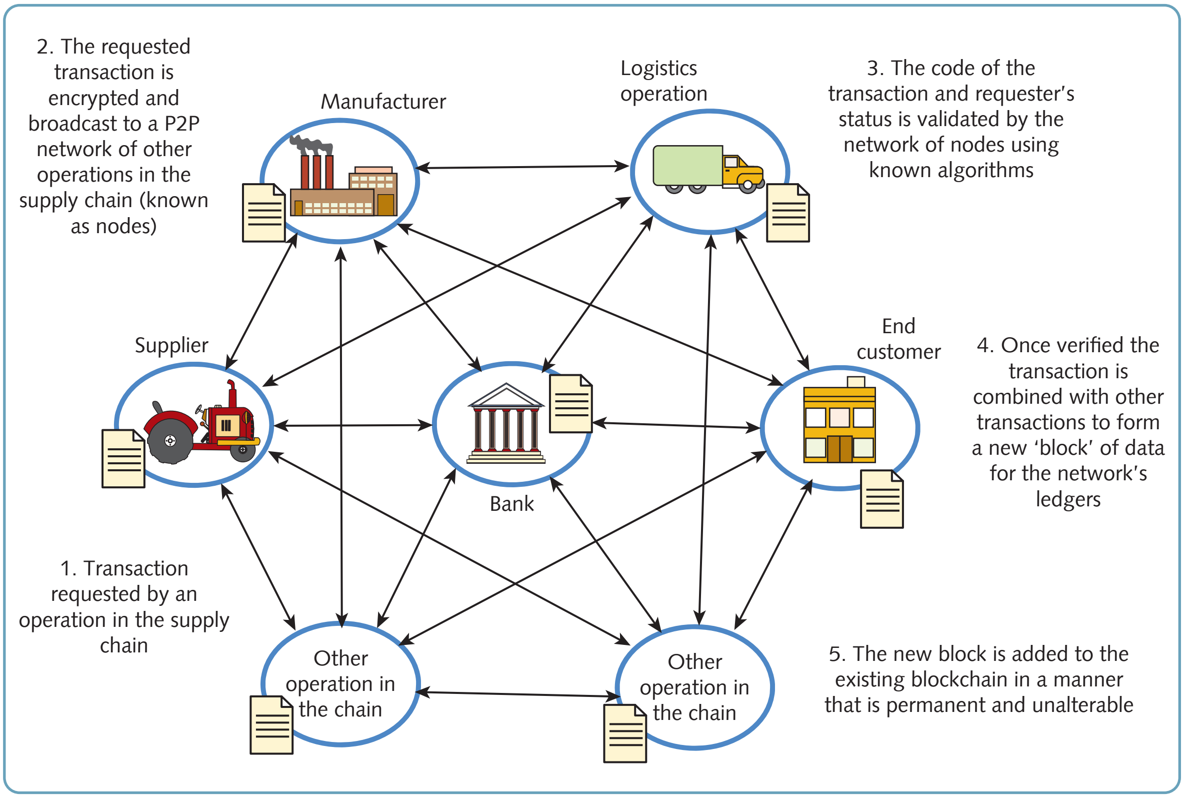


Figure 7.10 How the blockchain technology works to record and verify transactions in supply networks

Case example

TradeLens – blockchain revolutionises shipping[8]

Although the technology was still very much in its infancy, the global shipping giant Maersk announced in January 2018 that, subject to receipt of regulatory approvals, it was teaming up with IBM to form a joint venture to provide more efficient and secure methods for conducting global trade using blockchain technology. IBM was already recognised as a leading provider of blockchain technology. An early member of Hyperledger, an open source collaborative effort created to advance cross-industry blockchain technologies, it had worked with other clients to implement blockchain applications. The aim, it said, would be to offer a jointly developed platform, built on open standards and designed for use by the entire global shipping ecosystem. According to commentators, joining up with Maersk made sense because of the complexity of its integrated transport, container shipping, ports and logistics operations in 130 countries. More than 80 per cent (over $4 trillion) of the products consumers use daily are carried by the ocean shipping industry. But, traditionally, it had been a document-intensive business. The maximum cost of generating and processing all the required trade documentation to organise many of these shipments was estimated to be as high as one-fifth of the actual physical transportation costs. Blockchain technology was seen as ideally suited to large networks of disparate partners across complex supply chains. Maersk executives said that, with access to a shared, trusted record of transactions, the world's shipping companies would save money and be able to better compete on enhanced services.

The CEO of the joint venture, Michael J. White, said, *'Today, a vast amount of resources are wasted due to* *inefficient and error-prone manual processes. The pilots confirmed our expectations that, across the industry, there is considerable demand for efficiency gains and opportunities coming from streamlining and standardising information flows using digital solutions. Our ambition is to apply these learnings to establish a fully open platform whereby all players in the global supply chain can participate and extract significant value. We look forward to further expanding our ecosystem of partners as we progress toward a global solution.'*

By the end of 2018, Maersk and IBM announced the name of their platform venture – 'TradeLens'. Over the following years, the platform rapidly expanded in the shipping sector as many of the largest ocean carriers and their many ecosystem partners, representing more than 50 per cent of global container cargo, adopted TradeLens. The collaborative nature of the platform was critical in encouraging competitors, including CMA CGM, Hapag-Lloyd, MSC Mediterranean, Namsung and Ocean Network Express (ONE), to engage with TradeLens. For example, when announcing Hapag-Lloyd's adoption of the platform, Martin Gnass, Managing Director IT, said *'Expanding digital collaboration is critical to the evolution of the container shipping industry [. . .] we can collectively accelerate that transformation to provide greater trust, transparency and collaboration across supply chains and help promote global trade.'*

Jonathan NACKSTRAND/AFP/Getty Images

Critical commentary

- This emphasis on understanding the end customer in a supply chain has led some authorities to object to the very term, 'supply chain'. Rather, they say, they should be referred to as 'demand chains'. Their argument is based on the idea that the concept of 'supply' implies a 'push' mentality. Any emphasis on pushing products and services through a supply chain should be avoided. It implies that customers should consume what suppliers see fit to deliver. On the other hand, referring to 'demand chains' puts proper emphasis on the importance of seeing customers as pulling demand through the chain. Nevertheless, 'supply chain' is still the most commonly used term.

- When considering supply development, we focused on the most common approach, whereby a customer 'imposes' initiatives and performance requirements on its suppliers and sub-suppliers. Arguably, this perspective has two major limitations. First, most suppliers have several important customers that may all place conflicting development demands upon them. This can encourage suppliers to mislead customers over the extent to which they are adopting changes requested, which in turn may lead to trust erosion in trading relationships. Second, critics note that the traditional 'cascade' and 'intervention' approaches necessarily assume that the ideas flowing from the customer are superior to those of other parties in the supply network. However, in practice, many suppliers are equally, if not more, knowledgeable than their customers. These criticisms have led to some 'mutual' development initiatives whereby both customers and suppliers contribute knowledge and expertise in determining development priorities. Others have gone further, arguing that network development is more appropriate than supplier development. Here, the aim is to improve the overall network rather than a single organisation. Network development activities include quality assurance programmes, network-wide technology initiatives (see the case example on TradeLens), stakeholder suggestion programmes and supply network councils.

- The use of technology in supply chain management is not universally welcomed and can be viewed as preventing closer partnership relationships. In addition, many firms continue to underestimate the challenge of successfully assimilating such new technologies across their supply networks. As a result, the expected returns from some supply chain improvement initiatives fail to materialise. In addition, there are concerns that some supply chain technologies have the potential to threaten employment, individual privacy and supply chain security. For example, IoT could raise the potential for being hacked. Yet, any web-based connectivity will always create new vulnerabilities. In one view, we no longer have objects with computers embedded in them: we have computers with objects attached to them. Supply chain security is, arguably, given too little emphasis in the use of IoT.

SUMMARY CHECKLIST

- ☐ Is it understood that supply chain management (SCM) is equally concerned with flows of *information* as with flows of *products or services*?
- ☐ Is it understood that the performance of any one operation is partly a function of all the other operations in the supply network?
- ☐ Are supply chain concepts applied to internal (process-to-process) networks as well as external (operation-to-operation) networks?
- ☐ Are supply chain objectives understood in the context of the whole network rather than the single operation?
- ☐ Which product or service groups are ‘functional’ and which are ‘innovative’?
- ☐ Therefore, which products or service groups need efficient (‘lean’) and which need responsive (‘agile’) supply chain management?
- ☐ Is the position on the ‘transactional’ (or contractual) versus ‘partnership’ (or collaborative) spectrum understood for each supply chain relationship?
- ☐ Are relationships with suppliers and customers at an appropriate point on the transactional to partnership spectrum?
- ☐ Are ‘partnership’ relationships *really* partnerships or are they just called that?
- ☐ What approach to sourcing is most appropriate – multi sourcing, single sourcing, delegated sourcing or parallel sourcing?
- ☐ Has category analysis been carried out to determine the quadrant (non-critical, bottleneck, leverage or strategic) for different product or service purchases?
- ☐ Are suppliers selected using a rigorous and transparent process?
- ☐ Do those involved in negotiation understand different possible tactics (emotion, logic, threat, bargaining and compromise) and the potential risks of their use?
- ☐ Is there sufficient effort in ensuring a two-way flow of information between customers and suppliers?
- ☐ Are service-level agreements used to formalise customer expectations? And do they develop over time?
- ☐ Are mismatches between customer and supplier perceptions of both requirements and performance explored?
- ☐ Is sufficient effort put into supplier development?
- ☐ Are the logistics for any physical inputs and outputs managed effectively?
- ☐ Have technologies, such as the Internet of Things (IoT) or distributed ledger technology (DLT, also known as blockchain), been considered for adoption?
- ☐ What mechanisms for reducing the bullwhip effect have been implemented?

Case study

Big or small? EDF's sourcing dilemma

(This case was co-authored with Jas Kalra, Bartlett School of Construction and Project Management, University College London, and Jens Roehrich and Brian Squire, School of Management, University of Bath)

It was a warm afternoon as Stefano Moretto, Commercial Director of Hinkley Point C (HPC), and Eva Glines, Senior Supply Chain Engagement Manager, stood looking out of their office. Stefano, having recently joined EDF, had been tasked with establishing a supply network for the recently approved HPC project – the first of a number of new nuclear power stations aimed at supporting the United Kingdom (UK) Government's intended ambition of creating a 'clean, safe and affordable' energy future.

EDF and the Hinkley Point C (HPC) project

As with many other projects that the two had worked on, HPC would be complex and EDF would be right at the centre of things. The firm would hold detailed knowledge of the power station's design, maintain codes and standards, contribute to the design process at a strategic level and ensure the execution of the detailed design at an operational level. Over the coming years, EDF and its partners would need to build, test and commission two massive reactors, turbine halls and an electricity substation on the site, located in the South-West of the UK. Managing construction would involve people from many different organisations and disciplines working alongside each other. It would be EDF's responsibility to ensure that all relationships were underpinned by a consistent set of values and behaviours. However, at that moment, Stefano and Eva were preoccupied with another piece of the jigsaw – the site operations services.

Site operations services

Site operations services involved all the services not required for the actual construction of HPC. *People and organisation* services were there to make workers' lives more convenient and pleasant. Examples included catering, hospitality, cleaning, security and transportation. *Space and infrastructure* services were concerned with the physical infrastructure of the site. Examples included running the site's network of permanent and temporary roads safely and securely, as well as building maintenance, heating, lighting, plumbing, fire safety, etc. At the time, EDF expected to spend over £500 million on its site operations services over the period of HPC's construction (in fact, this figure rose to over £1 billion as the project developed). Given the large number of people likely to be on site at any one time, they were going to be vital to the project's success. '*It'll be like a small town*'. said Stefano. '*We're going to have a lot of people to move about, accommodate and feed.*' '*I know*', said Eva, '*and the clock's ticking! We need to make some decisions on the way we're going to go with this soon.*' The two then sat back down and began to discuss some of the key services needed at the HPC site and its associated development sites in nearby towns:

Catering

The HPC catering contract covered all aspects of catering at the HPC site and associated developments (AD). The scope of supply included: food production and preparation; food and beverage vending across the site; bar operations; management of food waste; chilled and ambient food distribution; cleaning and maintenance of kitchen and food servery areas; hospitality services; and management of all catering-related staff.

Transportation

A transport service provider would be responsible for operating and managing a bus service to transport construction and office workers between the HPC main site, associated development sites, car-parks and selected local towns in the area. Bus services would be operated in the following ways:

- *Park and ride:* Workers would embark at a single point of departure, travel directly to HPC main site, disembark at the site perimeter, pass through security and embark on an internal bus to be dropped off at the contractors' compound.
- *Direct service:* Buses would drop-off/pick-up at a number of different points to deliver workers to the HPC main site. They would then follow the same process as above.
- *Direct secure:* Operated from specific locations, workers would pass through security before embarking on buses that would deliver them directly to their place of work on site. These buses would pass through a 'fast-gate' at the site perimeter, negating the need for internal bus transfer.
- *Internal buses:* The transport provider would provide buses on the internal HPC site circuit to use in conjunction with the *park and ride* and *direct service* offerings.

Accommodation management

EDF would be building new campus accommodation buildings at the HPC site and in Bridgwater, a town close to the site. Once completed, the contracted firm would need to run these facilities. This would include: management of the hotel services for campus sites; day-to-day management of the catering contractor (see above); management of the campus buildings, campus grounds and sports facilities;

creating additional sales from hotel services; management of the security of guests while in the accommodation; 24/7 reception services; and management of all campus accommodation staff.

Facilities management

The facilities management contract would provide services to the HPC site and some of the associated developments. The scope of the services required would include: daily operational management of all temporary building facilities; general office services, including reception, porters and drivers, postal services and room booking; daily office cleaning, window cleaning and specialist cleaning (server rooms, etc.); domestic waste removal including confidential and segregated recycling; mechanical, electrical and building fabric maintenance, internal plumbing and drainage, and pest control; and audio-visual (AV) equipment management, maintenance and support.

Infrastructure operations and maintenance

EDF would construct the necessary permanent and temporary road network and other infrastructure for the HPC project both at the main construction site and at the associated development sites. Once constructed, these would need to be operated and maintained to ensure that the sites would be managed in a safe, secure and efficient way. The infrastructure operations and maintenance (O&M) contract would cover this activity.

Big or small?

'*OK, so we're pretty clear on what we need Eva. The next big question is big or small?*' Eva knew what Stefano was referring to – two competing views on the best strategy for sourcing EDF's site operations services. On one side was the argument that EDF should use one large 'generalist' supplier for each of the five main categories identified. These suppliers were typically multinational companies (MNCs) capable of providing a one-stop shop for the complete service solution required in any given contract. It was a tried-and-tested solution for sourcing in projects of this size. On the other side was the view that the firm should instead look to award contracts to local specialist small and medium-sized enterprises (SMEs) where possible. To do this would require breaking down some of the categories into more 'bitesize' contracts or possibly encouraging consortia of local suppliers who could jointly deliver site service requirements.

'*Eva, you know I've been working with some of the MNCs for years. They've got proven experience and expertise. Given that you, me and the team have got to set up over 150 tier-1 supply contracts over the next few months, maybe we're better going with what we know?*' Eva thought for a moment and responded: '*I see that argument, but sometimes I don't really rate these [MNC] firms. My experience is often they're hard work in negotiations and it doesn't get a whole lot better once they start providing the services. Also, they have too many other customers to be really concerned about giving us top service quality. At least smaller suppliers are likely to really put the effort in. Besides, I think we're agreed that it's at least worth thinking about whether we could really do something different this time*.'

'*All good points Eva*', said Stefano. He was also very keen to invite local SMEs to take advantage of the opportunities afforded by the HPC project. '*How good would it be to actually make a difference to the region through our supply chain?*', he thought. Still, earlier in the day, a conference call with another manager, discussing a leading global catering supplier intending to bid for the catering contract, had left Stefano unsure. His colleague had argued that it made more sense to let an experienced contractor with global presence help EDF manage the uncertainties associated with these contracts. Then he thought about some of the conversations that he'd had over the last two weeks. '*Another issue I'm thinking about here*,' Stefano continued, '*is what some of the service managers, who will eventually be responsible for the quality of services, have been saying to me. They seem to think that the local suppliers may not have the capability to deliver on this scale. Ultimately, if these services are not performed right, this project won't get off the ground*.'

Eva sighed. Stefano was right that most of the suppliers in the region were indeed small businesses, with no experience of delivering on the scale that would be required by EDF. But in the back of her mind was the feeling that not developing a local supply base would be a missed opportunity. '*I don't really agree with them on this Stefano. To be honest, much of this is just about a fear of doing things differently. Yes, MNCs are the safe and familiar option, but we're not talking about rocket science stuff here! All these services should be possible for local suppliers to deliver, surely?*'

Stefano thought for a moment. '*That's true Eva. Still, local SMEs are going to need a lot of upskilling to align with our needs. Developing a bespoke, local supply base with capabilities up to the quality standard we need will be a lot of effort. Remember, this project is already a very BIG jigsaw – I suspect some of the stakeholders might not think it's a good idea to add yet more pieces! And another thing, it's not just servicing the initial three or four thousand people on site that worries me, it's the ability to scale up to the seven, eight or nine thousand we'll probably have at the height of construction*.'

Eva shook her head. '*The services might be complex to manage, but they're not really capital- or technology-intensive. I think capability development in this area would be a lot easier to achieve than the manufacturing operations. From my conversations with the local chamber of commerce, it's clear that local businesses are very keen to work with us. We just need to develop a process of engaging more with local businesses, as I think quite of few of them wouldn't even think to bid for this work at the moment. I also wonder if we should encourage them to club together for some of these bits of work. It wouldn't just help now, but even more as we look to scale up further down the line*.'

Luke MacGregor/Bloomberg/Getty Images

'*Maybe Eva, maybe*.' Stefano quietly reflected that he was glad to be working with someone who was so passionate and shared his desire to do something value-added for the community. Eva's inputs certainly helped, but he remained undecided as to the best route forward. In addition to the issues they had just discussed, there was the broader political and public pressure to create economic and social value for the region. Construction of HPC was controversial and a recent piece of research had revealed that a substantial proportion of local residents were opposed to the project, while the acceptance of others was 'potentially fragile'. So, any good news stories were likely to be appreciated by the firm's leadership. Stefano stood up again. '*Come on Eva, I think we need to get the team together and try and make a decision. Let's go for a coffee first – I think we're in for a long day!*'

Note: All names and quantities mentioned in this case study are fictional.

Questions

1. How do the characteristics of different site operations services influence the sourcing decision (MNCs versus local SMEs)?
2. What other factors are affecting the decision?
3. If HPC were to adopt a sourcing strategy with a preference for local SMEs:
   (a) How might it engage effectively with local businesses to encourage bids?
   (b) How might it effectively configure these sourcing arrangements?
   (c) What approaches might it take to supplier capability development?

Appendix 1

Early expressions of interest for site operations service contracts

Catering	*Transportation*	*Accommodation management*	*Facilities management*	*Infrastructure operations and maintenance*
MNC1 Global, experienced caterer	**MNC2** International transport service provider	**MNC3** Global hotel construction and operations	**MNC4** Global facilities management and utility company	**MNC5** Global civil engineering contractor, highway, road construction, facilities maintenance
SME1 Wholesale ingredients and condiments	**NC1** National bus company	**SME8** Bed-and-breakfast owner	**SME11** Mechanical and electrical engineers	**SME13** Groundworks, civil engineering projects
SME2 Dairy and eggs producer	**SME7** Local bus company	**SME9** Independent laundry company	**SME12** Civil engineer and pipeline contractor	**SME14** Temporary traffic management service
SME3 Greengrocer		**SME10** Bricklaying, carpentry, decorating, plastering, rendering and roofing		
SME4 Tea and coffee supplier				
SME5 Baker				
SME6 Butcher				

Applying the principles

Some of these exercises can be answered by reading the chapter. Others will require some general knowledge of business activity and some might require an element of investigation. **Model answers for the first two questions are available on the student companion website**. *For model answers to the other questions in this section, please ask your tutor/lecturer.*

1. The COO of Super Cycles was considering her sourcing strategy. '*I have two key questions: for each of our outsourced parts, what is the risk in the supply market, and what is the criticality of the product or service to our business?*', Four key outsourced components are shown in Table 7.5. What approach to sourcing these components would you recommend?

2. A chain of women's apparel retailers had all their products made by Lopez Industries, a small but high-quality garment manufacturer. They worked on the basis of two seasons: Spring/Summer and Autumn/Winter. '*Sometimes we are left with surplus items because our designers have just got it wrong*', said the retailer's chief designer. '*It is important that we are able to flex our order quantities from Lopez during the season. Although they are a great supplier in many ways, they can't change their production plans at short notice*.' Lopez Industries was aware of this. '*I know that they are happy with our ability to make even the most complex designs to high levels of quality. I also know that they would like us to be more flexible in changing our volumes and delivery schedules. I admit that we could be more flexible within the season. Partly, we can't do this because we have to buy-in cloth at the beginning of the season based on the forecast volumes from our customers. Even if we could change our production schedules, we could not get extra deliveries of cloth. We only deal with high-quality and innovative cloth manufacturers who are very large compared to us, so we do not represent much business for them*.' A typical cloth supplier said: '*We compete primarily on quality and innovation. Designing cloth is as much of a fashion business as designing the clothes into which it is made. Our cloth goes to tens of thousands of customers around the world. These vary considerably in their requirements, but presumably all of them value our quality and innovation*.' How should the retailers try to influence this supply chain in order to improve its performance?

3. The example of the bullwhip effect shown in Table 7.4 shows how a simple 5 per cent reduction in demand at the end of a supply chain causes fluctuations that increase in severity the further back an operation is placed in the chain.

   a) Using the same logic and the same rules (i.e. all operations keep one period's inventory), what would the effect on the chain be if demand fluctuated period by period between

Table 7.5 Four outsourced components for Super Cycles

Component	*Cost (as a proportion of total material cost)*	*Suppliers*	*Ease of changing supplier*
The inner tubes	3%	Many alternative suppliers	Very easy; could do it in days
Frame tubing	15%	Only one supplier capable at the moment; could take a long time to develop a new supplier	Difficult in the short term, possible in the longer term
Carbon-fibre stem and bars	32%	Relatively large number of available suppliers	Relatively easy; would probably take a few weeks for new contract
'Groupset' gearing system	35%	Few suppliers who are capable of manufacturing these components to sufficient quality	Complex to source; could switch supplier in the longer term but would pose quality risk

100 and 95? That is, period 1 has a demand of 100, period 2 has a demand of 95, period 3 a demand of 100, period 4 a demand of 95, and so on?

b) What happens if all operations in the supply chain decide to keep only half of the period's demand as inventory?

c) Find examples of how supply chains try to reduce this bullwhip effect.

4. If you were the owner of a small local retail shop, what criteria would you use to select suppliers for the products that you wish to stock in your shop? Visit three shops that are local to you and ask the owners how they select their suppliers. In what way were their answers different from what you thought they might be?

5. Many companies devise a policy on ethical sourcing, covering such things as workplace standards and business practices, health and safety conditions, human rights, legal systems, child labour, disciplinary practices, wages and benefits, etc.

a) What do you think motivates a company to draw up a policy of this type?

b) What other issues would you include in such a supplier selection policy?

6. Airline catering is a tough business. Meals must be of a quality that is appropriate for the class and type of flight, yet the airlines who are its customers are always looking to keep costs as low as possible. Menus must change frequently and respond promptly to customer feedback. Forecasting passenger numbers is difficult. Suppliers are advised of likely numbers for each flight several days in advance, but the actual minimum number of passengers is only fixed six hours before take-off. Also, flight arrivals can be delayed, upsetting work schedules – even when on time, no more than 40 minutes are allowed before the flight takes off again. Airline caterers usually produce food on, or near, airports using their own staff. Catering companies' suppliers are also usually airline specialists that are also located near the caterers. A consortium of Northern Foods, a leading food producer (that normally supplies retailers), and DHL won a large contract at Heathrow Airport against the traditional suppliers. DHL was already a large supplier to 'airside' caterers there, with its own premises at the airport. Northern Foods made the food at its existing factories and delivered it to DHL, where it was assembled onto airline catering trays and transferred to the aircraft.

a) Why would an airline use a catering services company rather than organise its own on-board services?

b) What are the main operations objectives that a catering services company must achieve in order to satisfy its customers?

c) Why is it important for airlines to reduce turnaround time when an aircraft lands?

d) Why was the Northern Foods–DHL consortium a threat to more traditional catering companies?

Notes on chapter

1 The information on which this example is based was taken from: Banker, S. (2017) 'Drones deliver life-saving supplies in Africa', *Forbes*, 13 October; Stewart, J. (2017) 'Blood-carrying, life-saving drones take off for Tanzania', *Wired*, 24 August; flyzipline.com/how-it-works/ [accessed 18 September 2020].

2 Fisher, M.L. (1997) 'What is the right supply chain for your product?', *Harvard Business Review*, 75 (2).

3 The information on which this example is based was taken from: author visit to The Donkey Sanctuary, Sidmouth, UK, 2020, https://www.thedonkeysanctuary.org.uk/; Hameed, A., Tariq, M. and Yasin, M.A. (2016) 'Assessment of welfare of working donkeys and mules using health and behavior parameters', *Journal of Agricultural Science and Food Technology*, 2 (5), pp. 69–74.

4 This approach is usually credited to Kraljic; it is often referred to as the Kraljic Matrix, see Kraljic, P. (1983) 'Purchasing must become supply management', *Harvard Business Review*, September.

5 Our thanks to Professor Mark Johnson of Warwick Business School for this example.

6 Thanks to Stephen Disney at University of Exeter for help with this section.

7 Iansiti, M. and Lakhani, K.R. (2017) 'The truth about blockchain', *Harvard Business Review*, January–February.

8 The information on which this example is based was taken from: del Castillo, M. (2018) 'Shipping Blockchain: Maersk spin-off aims to commercialize trade platform', Coindesk.com, 16 January; Slocum, H. (2018) 'Maersk and IBM to form joint venture applying blockchain to improve global trade and digitize supply chains', press release, Maersk, 18 January; Maersk (2019) 'TradeLens blockchain-enabled digital shipping platform continues expansion with addition of major ocean carriers Hapag-Lloyd and Ocean Network Express', press release, Maersk, 2 July.

Taking it further

Chopra, S. and Meindl, P. ***(2015) Supply Chain Management: Strategy, planning, and operation, 5th edition, Pearson.*** *One of the best of the specialist texts.*

Christopher, M. ***(2016) Logistics and Supply Chain Management, 5th edition, Pearson.*** *Updated version of a classic that gives a comprehensive treatment of supply chain management by one of the gurus of the subject.*

Voss, C. and Raz, T. ***(2016) Never Split the Difference: Negotiating as if your life depended on it, Random House.*** *A really interesting book examining a wide range of negotiation settings and building on some of the ideas of mutual gains bargaining and win–win approaches to negotiation.*

Vyas, N., Beije, A. and Krishnamachari, B. ***(2019) Blockchain and the Supply Chain: Concepts, strategies and practical applications, Kogan Page.*** *A practical book examining the potentially transformative effect of blockchain technology on supply networks, plus some very interesting insights on the origins of supply chain management, starting with Egypt in 2500BC!*

8 Capacity management

Introduction

Capacity management is the activity of coping with mismatches between the nature of an operation's demand and its ability to supply. It involves an attempt to forecast demand and measure the ability to supply products and services, followed by the selection of appropriate demand-side and supply-side responses based on performance objectives and long-term outlook. In doing this, operations managers must reconcile two competing requirements. On the one hand is the importance of maintaining customer satisfaction by delivering products and services to customers reasonably quickly. On the other is the need for operations (and their extended supply networks) to maintain efficiency by minimising the costs of excess capacity. This is why capacity management is so important – it has an impact on both revenue and costs, and therefore profitability (or the general effectiveness of service delivery in not-for-profit operations). In this chapter, we look primarily at these competing tensions at an aggregated and medium-term level. Figure 8.1 shows where capacity management fits into the overall structure of the text.

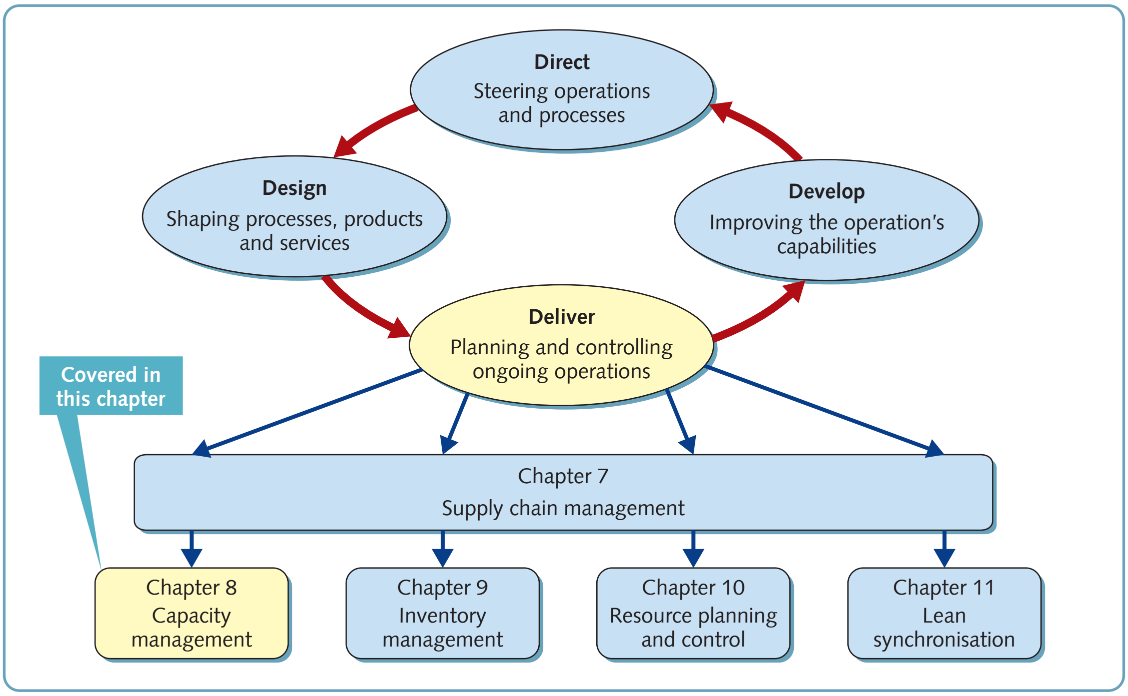


Figure 8.1 Capacity management is the activity of coping with mismatches between demand and the ability to supply demand

EXECUTIVE SUMMARY

8.1 Is the importance of capacity management understood?

↓

8.2 Have patterns of demand been understood and influenced?

↓

8.3 Has the operation's current capacity been determined?

↓

8.4 Have strategies for managing the 'supply side' been considered?

↓

8.5 Have the consequences of capacity management decisions been considered?

8.1 Is the importance of capacity management understood?

Capacity management is the activity of understanding the nature of demand and supply, and of selecting alternative demand-side and supply-side strategies to meet the needs of customers. In the long term, capacity decisions include total capacity, its distribution between sites and its location. In the medium term, capacity decisions include the extent to which capacity levels fluctuate, staffing levels and the degree of subcontracting. In the short term, capacity decisions include which resources should be allocated to each set task, and how tasks should be allocated to resources. Longer time frame capacity decisions impose constraints on shorter time frame decisions, while decisions made for shorter time frames provide important input into longer-term decisions.

8.2 Have patterns of demand been understood and influenced?

There are two key tasks on the demand side of the capacity management framework. The first is to understand patterns of demand over different time frames (e.g. hourly, daily, weekly, monthly, annually, etc). A number of qualitative approaches can be used to support this analysis including panel, Delphi and scenario planning. Quantitative approaches include time-series methods (moving-average, exponential smoothing and the multiplicative seasonal model) and causal (associative) methods. The second task is to decide if and how to influence patterns of demand for products and services. Demand management methods include constraining customer access, price differentials, scheduling promotions, service differentials and creating alternative products and services. When operations have relatively fixed capacities, yield management is often used to maximise revenues.

8.3 Has the operation's current capacity been determined?

Every operation needs to know their capacity, because if they have too little, they cannot meet demand (bad for revenue) and if they have too much, they are paying for more capacity than they need (bad for costs). However, measuring capacity can be relatively difficult unless the operation is standardised and repetitive. Capacity estimates depend on the mix of products or services supplied – as the mix changes, so too does the effective capacity. They also depend on the time over which they are supplied – the level of activity and output that may be achievable over short periods of time is not the same as the capacity that is sustainable on a regular basis. Specification of products and services also plays a role in capacity calculations, with some operations able to increase their output by reducing the specification of their output during busy periods. A reduction in capacity is sometimes called 'capacity leakage', one method of measuring which is called overall equipment effectiveness (OEE).

8.4 Have strategies for managing the 'supply side' been considered?

Having determined the capacity of the operation, managers must make a number of important supply-side decisions. The first is setting 'base capacity', which is influenced by three main factors: the relative importance of the operation's performance objectives; the perishability of the operation's outputs; and the degree of variability in demand or supply. High service levels, high perishability of an operation's outputs and a high degree of variability either in demand or supply, all encourage a relatively high level of base capacity. Managers must then evaluate two alternative methods of managing supply. 'Level capacity' plans involve no change in capacity and require that the operation absorb demand–capacity mismatches, usually through under- or overutilisation of its resources, or the use of inventory. 'Chase capacity' plans involve adjusting capacity through methods such as overtime, varying the size of the workforce or subcontracting.

8.5 Have the consequences of capacity management decisions been considered?

When making capacity management decisions, firms look to balance the need for customer service with the need to maintain operational efficiency. In practice most use a combination of demand-side and supply-side methods to achieve this balance. There are three useful ways of assessing the mix of 'pure' capacity management plans (demand management, level capacity or chase capacity) that may be most appropriate for an organisation. First, managers should analyse the relative balance of predictable versus unpredictable demand variation, given its effect on objectives and tasks of capacity management. Second, when outputs can be stored, cumulative representations of demand and capacity are a useful method for planning capacity. Third, taking a longitudinal perspective to consider short- and long-term outlooks helps align different time frames for capacity investments (or divestments). Finally, capacity management is a dynamic process with decisions reviewed period by period. It is essential that capacity decisions made in one period reflect the knowledge accumulated from experiences in previous periods.

8.1 Diagnostic question: *Is the importance of capacity management understood?*

Capacity management is concerned with understanding the nature of demand and supply (capacity), and attempting to reduce mismatches between them. Demand is the quantity of products or services that customers request from an operation or process at any point in time. Capacity is the output that an operation or process can deliver over a defined unit of time. This 'ability-to-supply' perspective on capacity is a broad view in that it depends on *all* the stages in a given process, operation or supply network. So, for example, the capacity of an ice cream manufacturer is determined by how much ice cream its factories can produce in a given period of time, but also by how much packaging material, raw material supplies, and so on, that its suppliers can provide. It may have the factories to make 10,000 kilos of ice cream a day, but if its suppliers of dairy products can only supply 7,000 kilos a day, then the effective capacity (in terms of 'ability to supply') is limited to 7,000 kilos per day. Of course, if demand remains steady, any operation will attempt to make sure that its supply capacity does not limit its own ability to supply. However, given that capacity management is concerned with fluctuations in demand *and* supply, the process of managing capacity is essentially a dynamic issue.

> OPERATIONS PRINCIPLE
> ***Any measure of capacity should reflect the ability of an operation or process to supply demand.***

A mismatch between demand and capacity can occur because demand fluctuates over time, capacity fluctuates over time, or both. It is worth noting that 'coping' with mismatches between demand and capacity does not always mean that capacity should match demand. Many operations take the deliberate decision to fail to meet demand or fail to fully exploit its ability to supply. For example, a hotel may not make any effort to meet demand in peak periods because doing so would incur unwarranted capital costs. It is therefore content to leave some demand unsatisfied, although it may increase its prices to reflect this. Similarly, a flower grower may not supply the entirety of its potential supply (crop) if doing so would simply depress market prices and reduce its total revenue.

Integrating across the levels of capacity management

The central issue of capacity – deciding if and how to meet demand – can be considered at different organisational levels and over different time periods. In Chapter 4, we examined long-term capacity decisions as they relate to the structure and scope of operations. At this strategic level, capacity decisions include: How much capacity do we need in total? How should the capacity be distributed? Where should the capacity be located? In this chapter, we focus more on the medium-term aspects of capacity management, where decisions are being made largely within the constraints of the physical capacity set by the operation's long-term structure and scope. Medium-term capacity management usually involves assessing demand forecasts with a time horizon of between 2 and 18 months, during which time planned output can be varied, for example, by changing the number of hours that resources are used. In practice, however, few forecasts are accurate, and most operations also need to respond to changes in demand that occur over even shorter timescales. We examine these 'short-term' decisions, around task allocation, sequencing and resourcing, in Chapter 10.

While, for practicality, we have divided long-term, medium-term and short-term aspects of capacity management in this text, it is important to note the interaction effects between different time horizons (illustrated in Figure 8.2). Capacity decisions are made within the constraints of the physical limits of the operation, the ability of its suppliers to supply, the availability of staff, and so on. In the other direction, shorter-term decisions provide important feedback for planning over longer-term time horizons.

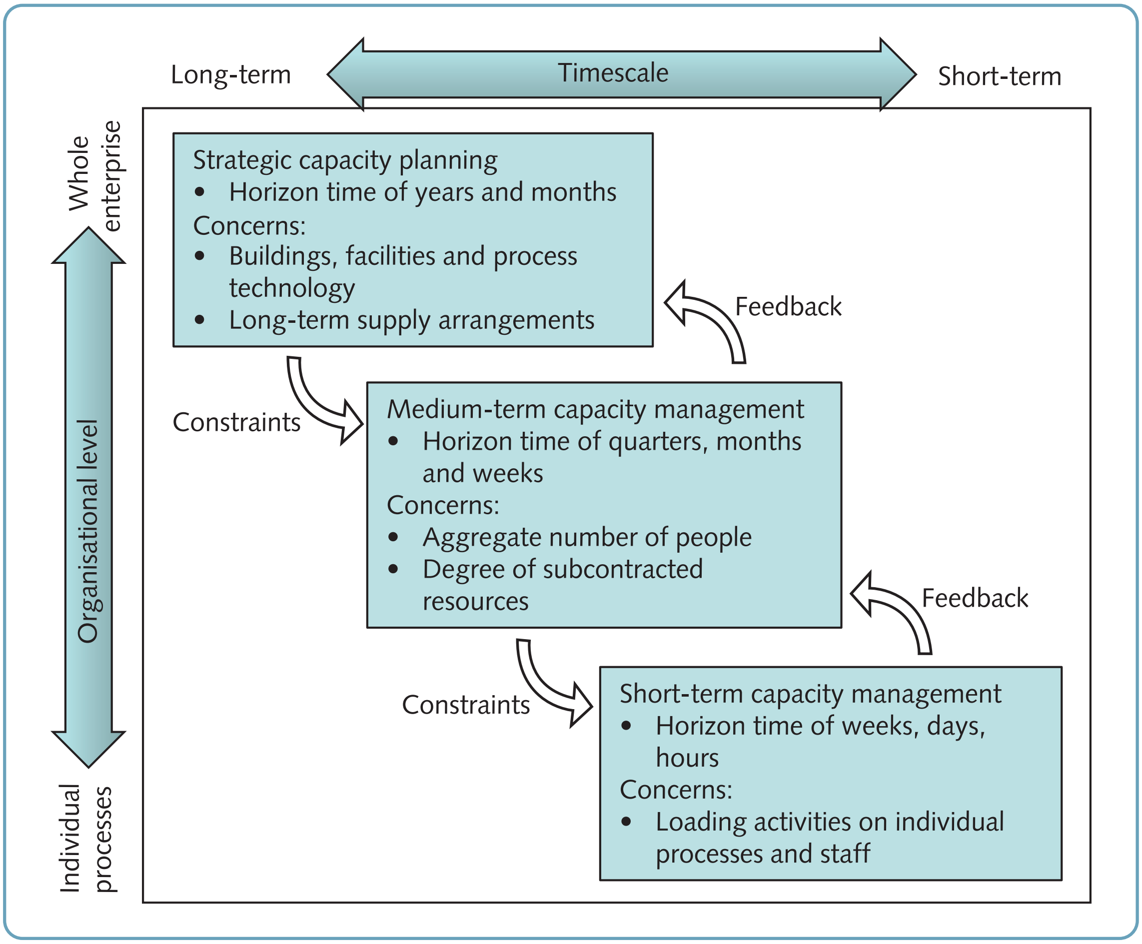


Figure 8.2 **Capacity management should be integrated across levels because each level constrains what can be done in the level below and can provide feedback for the level above**

Case example

The rise of the gig and sharing economies[1]

At one time, most people had one more-or-less permanent job for life, and their possessions were for their exclusive use. But then things changed, and the change had a profound effect on capacity management. What changed was the rise of the so-called 'gig economy' and the 'sharing economy' – two related but different phenomena. The gig economy describes the trend for organisations to employ subcontractors on a freelance basis rather than relying on full-time employees. The sharing economy describes how individuals or groups are able to hire out any underutilised assets that they own, often on a short-term basis. As a method of flexing capacity, both are employed across a wide range of industries, including arts and design, transportation, construction, accommodation, media, education and professional services. Here are three examples of companies at the forefront of the gig and sharing revolutions.

Uber

Uber has become, arguably, the most famous of all gig economy companies worldwide. Its technology platform connects those wanting taxi rides, food delivery and transportation of small packages, with individuals who want to provide a service. It gives subcontracted drivers considerable flexibility over when and where they work, generating what has been described as a perfectly competitive supply market. Uber has also developed other operations in over 600 cities worldwide with extensions to its core ride-sharing service, including Uber Boat (a water-taxi service), Uber-MOTO (transportation by motorcycle), Uber Eats (a meal delivery service) and UberRUSH (a courier package service).

Udemy

The online education provider Udemy's peer-to-peer platform hosts more than 150,000 courses developed by around 60,000 independent instructors to a student base of over 50 million. Courses cover a wide variety of subjects including business, arts, health and fitness, languages and technology. Given the sheer scale of its user base, online content developers are able to generate good returns on the time they have invested – some are millionaires. Similar companies include Coursera, Kahn Institute, Canvas and FutureLearn, all enabling freelance experts to post content. Some, like Udemy, operate as for-profit, while others are not-for-profit platforms looking to 'democratise learning'.

Airbnb

Airbnb provide a peer-to-peer platform connecting customers with those wanting to lease lodgings, including houses, apartments, bedrooms, hotel rooms and hostels. They generate income by taking a commission on every transaction on their market-place. Airbnb's millions of listings worldwide facilitate over 300 million check-ins. In its efforts to increase trust between customers and suppliers of capacity, Airbnb use a rating system and Airbnb Plus, which features only homes vetted for quality and comfort. Those not wanting the hassle of preparing their property might instead choose OneFineStay, a competitor that, in addition to undertaking quality vetting, caters to customers by providing greeting services, key handling, insurance, cleaning and provision of bed linen.

Other companies that have emerged using platforms acting as intermediaries between demand (customers) and supply include those for leasing cars, vans and even aircraft, website development, graphic design, finance and consulting advice, programming and chef services. From a capacity management perspective, these developments help to maintain high levels of customer service, even in the face of changeable demand, while also achieving high levels of utilisation. Operations avoid the fixed costs of employees or facilities when demand drops but can quickly ramp up capacity when demand increases. However, 'gig working' is seen by some as unethical, and in some countries moves towards greater legal rights for 'gig workers' (such as sick pay, maternity leave and paid holidays) have reduced some of its operational advantages for employers.

A framework for capacity management

The activities involved in capacity management are shown in Figure 8.3. On the demand side of the framework, the first step is to measure demand for products and services. This involves selecting from a range of qualitative and quantitative tools to support more accurate forecasting. The next step is to consider if and how to change demand patterns using demand management or yield management. On the supply side of the framework, the first step involves

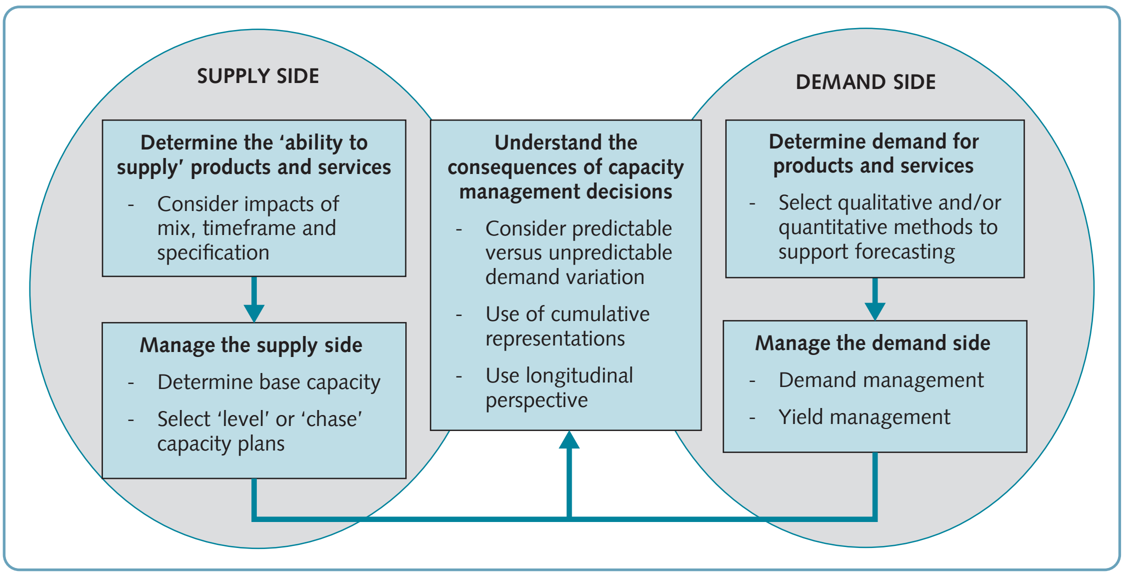


Figure 8.3 Capacity management framework

measuring the 'ability to supply' products and services. Here, the impacts of product or service mix, time frame and output specification should be considered. The next step is to manage the supply side by determining the appropriate 'base level' of capacity and then deciding whether to keep this constant over time (level capacity plan) or make adjustments in line with changing demand patterns (chase capacity plan). Finally, managers must understand the consequences of different capacity management decisions. This involves considering the mix of predictable and unpredictable demand variation, the use of cumulative representations and taking a perspective that combines both short-term and long-term outlooks.

8.2 Diagnostic question: *Have patterns of demand been understood and influenced?*

On the demand side of the capacity management framework, there are two related tasks: (a) understanding patterns of demand; and (b) deciding if and how to influence these patterns. Managers need to consider demand patterns over different time frames (e.g. hourly, daily, weekly, monthly, annually, etc), and various qualitative and quantitative methods are useful in improving forecasts. In this chapter, we briefly outline some of the most common methods, while further details can be found in the forecasting supplement at the end of Chapter 4. In seeking to then influence demand patterns, a range of demand management and yield management tools may be applied.

> OPERATIONS PRINCIPLE
> ***Understanding patterns of demand for products and services is critical for successful capacity management.***

Although demand forecasting is often the responsibility of the sales and/or marketing functions, it is a very important input into the capacity management decisions, and so is of interest to operations managers. After all, without an estimate of future demand it is not possible to plan effectively for future events; one can only to react to them as they happen.

Regardless of context, operations should consider a number of key questions when considering demand:

- What is the overall demand for a product or service over a period of time?
- In any period of time, how much does demand change?
- Are the changes in demand easy or difficult to predict?
- Can demand patterns be influenced, and if so, what are the most suitable methods to adopt?

Approaches to demand forecasting

In the supplement to Chapter 4, we described two core approaches to forecasting – qualitative and quantitative. Qualitative techniques rely on perception, instinct, expert opinion and consensus building. In contrast, quantitative techniques predict demand mathematically, using historical data and causal variables. Both approaches have their strengths and limitations. As such, arriving at higher-quality predictions of future demand often requires the use of qualitative and quantitative methods in combination.

Qualitative approaches to forecasting

There are a variety of qualitative methods used by organisations looking to make predictions about the future demand for products or services. These include:

- *Panel approach* – where a number of experts (instead of a single person) attempt to reach a consensus forecast.

- *Delphi method* – a specific form of panel approach, where experts make initial estimates and review an anonymous summary of estimates from other experts on the panel, after which their original estimates can be revised – a process that may be repeated several times.
- *Scenario planning* – a method of long-term forecasting, where experts develop and then discuss the operational implications for a range of future scenarios.

Quantitative approaches to forecasting

Some organisations prefer to use a quantitative approach to forecast demand. There are a number of methods used, including:

- *Time series* – a set of methods including moving-average, weighted-moving-average, exponential smoothing and multiplicative seasonal models, which use historical data to model future demand patterns.
- *Causal (associative) models* – these typically rely on regression analysis or structured equation modelling to describe and evaluate the complex cause–effect relationships.

Case example

How artificial intelligence helps with demand forecasting[2]

Of the many potential applications of artificial intelligence (AI), arguably one of the most useful is its use in demand forecasting. In particular, advances in machine learning have improved the availability and reliability of forecasts as well as the speed with which they can be produced. Like traditional methods of forecasting, AI-based systems begin with a set of assumptions and rules that, when applied to available data, produce a forecast. The 'deep learning' of AI methods can learn and adapt these rules using huge amounts of historical data. However, machine learning models suffer from one major limitation – they work far more effectively when input data is similar to the data upon which they were 'trained'. In the face of sudden unexpected rises or falls in demand, predictive algorithms struggle to reflect the reality on the ground, because the system has not been 'trained' for any particularly large changes. In other words, they become unreliable when the past is incapable of giving helpful insights into the future.

For example, the COVID-19 pandemic had impacts on almost all facets of business life and one of the most dramatic areas was the effect on demand patterns. Around the world, firms manufacturing face masks, hand sanitisers, cleaning products, toilet paper, home and garden equipment, medical supplies, canned and frozen goods, games and home fitness apparatus, saw dramatic increases in demand over short time periods. Likewise, the demand soared for services such as food delivery, logistics, online video calling and social media platforms. Other firms experienced massive drops in their demand – those connected with the travel, oil and gas, hospitality and leisure, and auto industries were particularly affected. The consequence on demand forecasting was significant. Many firms experienced an almost overnight shift from relatively stable and predictable demand variation – the easiest pattern for an operations manager to deal with, to much more volatile and less predictable demand variation – the most challenging pattern to deal with. This meant that much of the historical data that would normally be essential in developing high-quality demand forecasts were instead wildly inaccurate. And for those firms who had invested in machine learning models to support demand forecasting, the artificial intelligence (AI) algorithms were simply incapable of dealing with the radical shifts in 'normal' human behaviour during the crisis.

This fundamental limitation meant that many firms required significant manual intervention to override

chombosan/Alamy Stock Photo

automatically generated capacity plans and left some commentators questioning the suitability of machine learning models (and indeed other methods relying on historical data), especially in more dynamic business contexts. However, the counter argument was that the nature of AI is a 'learning technology'. With the right kind of human intervention, machine learning models could be 'retrained' to take account of more dynamic data sets. Some experts have suggested that one approach would be for AI training to incorporate much longer time frames and, in doing so, include other 'black swan' events over the past century. In this way, they argue, the accuracy of machine learning models could be improved significantly.

Making forecasts useful for capacity management

Without some attempt to understand future demand fluctuations (and indeed supply fluctuations – see later in the chapter), it is not possible to plan effectively for future events, only to react to them. There are three ways to assess the usefulness of a demand forecast from an operations manager's perspective – its level of accuracy, its ability to indicate relative uncertainty, and its expression in terms that are useful for capacity management.

Demand forecasts should be as accurate as possible

The process of capacity management is hugely aided if forecasts are as accurate as possible. Firms must often decide output in advance, based on a forecast that might change before the demand occurs or, worse, prove not to reflect actual demand at all. So, while no approach will result in a perfect forecast, a combination of qualitative and quantitative approaches can be used to great effect by bringing together expert judgements and predictive models. In addition, managers need to be aware that some forecast errors are particularly damaging. For example, if a process is operating at a level close to its maximum capacity, over-optimistic forecasts could result in unnecessary capital expenditure to increase capacity. In contrast, inaccurate forecasts for an operation well below its capacity limit will result in extra cost, but probably not to the same extent. So, critically, the effort put into forecasting should reflect the varying sensitivity to forecast error.

Demand forecasts should give an indication of relative uncertainty

Demand in some periods is more uncertain than others. The importance of this is that the operations managers need an understanding of when increased uncertainty makes it necessary to have reserve capacity. A probabilistic forecast allows this type of judgement, between plans that would virtually guarantee the operation's ability to meet actual demand and plans that minimise costs. Ideally, this judgement should be influenced by the way the business wins orders: price-sensitive markets may require a risk-avoiding cost minimisation plan that does not always satisfy peak demand, whereas markets that value responsiveness and service quality may justify a more generous provision of operational capacity.

Demand forecasts should be expressed in terms that are useful to capacity management

If forecasts are expressed only in monetary terms and give no indication of the demands that will be placed on an operation's capacity, they will need to be translated into realistic expectations of demand, expressed in the same units as the capacity. Nor should forecasts be expressed in monetary terms, such as sales, when those sales are themselves a consequence of capacity planning. For example, some retail operations use sales forecasts to allocate staff hours throughout the day. Yet sales will also be a *function* of staff allocation. Better to use forecasts of 'traffic' – the number of customers who will potentially want to be served if there are sufficient staff to serve them.

Demand management

> OPERATIONS PRINCIPLE
> ***Demand management involves changing the pattern of demand by stimulating off-peak demand or constraining peak demand.***

Demand patterns clearly have a big influence on the way operations function. For this reason, many organisations are not content to simply attempt to improve their demand forecasts. Instead they seek to change the patterns of demand in some way to bring them closer to available capacity. Figure 8.4 illustrates how this is achieved – by either stimulating off-peak demand or by constraining peak demand. There are a number of methods used in demand management:

- *Constraining customer access* – customers may only be allowed access to the operation's products or services at particular times. For example, reservation and appointment systems in various service contexts.
- *Price differentials* – adjusting price to influence demand by increasing prices during high demand periods and reducing prices during low demand periods.
- *Scheduling promotion* – varying the degree of market stimulation through promotion and advertising in order to encourage demand during normally quieter periods.
- *Service differentials* – adjusting (implicitly or explicitly) service levels to reflect demand by allowing service to deteriorate in periods of high demand and increase in periods of low demand. If this strategy is used explicitly, customers are being educated to expect varying levels of service and hopefully move to periods of lower demand.
- *Creating alternative products and services* – attempting to create alternative products or services to fill capacity in quiet periods. For example, ski resorts may provide organised mountain activity holidays in the summer. However, the apparent benefits of filling capacity in this way must be weighed against the risks of damaging the core offering, and the operation must be capable of serving both markets effectively.

Yield management

One approach used to maximise revenues by operations that have relatively fixed capacities is yield management. This approach is especially useful when the market can be fairly clearly segmented, the service cannot be stored in any way, the service is sold in advance and the marginal cost of making a sale is relatively low. Airlines, for example, fit all these criteria. They adopt a collection of methods to try to maximise the yield (i.e. profit) from their capacity. Overbooking capacity may be used to compensate for passengers who do not show up for the flight. However, if more passengers show up than they expect, the airline will have a number of upset passengers (although they may be able to offer financial inducements for the passengers to take another flight). By studying past data on flight demand, airlines try to balance the risks

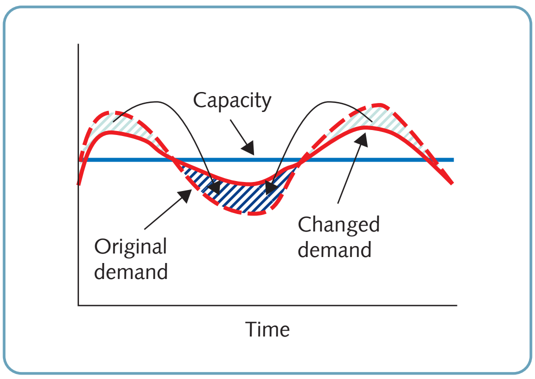


Figure 8.4 Demand management plan

Case example

Demand management using surge pricing for cab hailing and art galleries[3]

Surge (or dynamic) pricing is a demand management technique that relies on frequent adjustments in price to influence supply and (especially) demand. For example, some electricity suppliers charge different rates for energy depending on when it is consumed. Similarly, in countries with road charging, tolls are set at higher levels during peak times in an effort to keep traffic flowing. But perhaps the best-known example of surge pricing is the algorithm used by the cab-hailing app Uber. During times of excessive demand or inadequate supply, when the number of people wanting a ride exceeds the number of available drivers, Uber applies a multiplier to increase its normal fares based on the scarcity of available drivers. The company says that it does this to make sure those who need a ride can get one. Moreover, surge pricing helps ensure that pick-ups are available quickly and reliably. For its drivers, surge means higher fares and a steady stream of ride requests.

Even leading arts venues in London, such as the Royal Academy of Arts, the Barbican and the National Portrait Gallery, introduce 'surge pricing', by charging higher prices at popular times. The justification used by arts venues is usually that the extra income helps them to subsidise admissions for less popular exhibitions. It also helps to smooth visitor numbers throughout the week. However, some critics dislike the intrusion of such commercial demand-management methods into the art world. Even in the unambiguously commercial world of cab hailing, surge pricing can be deeply unpopular with customers. In the press and on social media, customers complain that Uber are taking advantage of them. But some marketing experts say it is, at least partly, a matter of perception, and as well as capping their multiplier, Uber should make the way it calculates it more transparent, limit how often prices are adjusted, communicate the benefits of the technique, and change its name (*certainty pricing* and *priority pricing* have been suggested).

metamorworks/iStock/Getty Images Plus/Getty Images

of overbooking and underbooking. Operations may also use price discounting at quiet times, when demand is unlikely to fill capacity. For example, hotels will typically offer cheaper room rates outside of holiday periods to try to increase naturally lower demand. In addition, many larger hotel chains will sell heavily discounted rooms to third parties who in turn take on the risk (and reward) of finding customers for these rooms.

8.3 Diagnostic question: *Has the operation's current capacity been determined?*

We now turn to the supply side of the capacity management framework. Here, the first task is to determine the capacity (i.e. ability to supply) of the operation. Capacity is the *maximum level of value-added activity over a period of time* that the process can achieve under normal operating conditions. Critically, this definition reflects the scale of capacity but, more importantly, its *processing capabilities*. Measuring these processing capabilities may sound simple but can in fact be relatively hard to define unambiguously unless the operation is standardised and repetitive. Any measure of capacity will contain a number of assumptions, each of which may be necessary to give an estimate of capacity, but each of which obscures some aspect of reality. Again,

> OPERATIONS PRINCIPLE
> ***Measuring capacity is a function of product/service mix, duration and product/service specification.***

taking capacity as 'the ability to supply', these assumptions relate to the mix of products or services supplied, the time over which they are supplied and the specification of what is supplied.

Case example

Next-generation signal technology expands railway capacity[4]

In many parts of the world, railway networks are at breaking point, struggling to deal with the demands placed on them daily. At peak times of the day, some commuter lines are near gridlock, and many commentators put the blame on capacity shortages. But recall our earlier definition – the capacity of an operation is the *maximum level of value-added activity over a period of time* that can be achieved under normal operating conditions. Therefore, for railway networks, as with airports, seaports, roads and other transport infrastructure, it's not just about what resources you have (the scale of capacity), it's about what you can do with it (the processing capabilities of capacity). Herein lies the problem – increasing the scale of capacity is often difficult. It can be expensive, very slow and often politically sensitive. Therefore, while they wait for longer-term increments in capacity, railways (as with other operations) often have to make the best of what they have.

One way this can be achieved is through the digitisation of railway signalling systems. Traditional signalling was developed over a century ago using an approach called 'block working' whereby only one train can be in a 'block' at any one time. However, the fixed position of signals means that it is impossible to increase capacity without risking safety. In addition, many railways around the world now combine a wide array of different systems to manage trains – in one area there may be advanced software automatically routing train pathways, in another signal workers continue to operate manual wire-systems to control semaphore arms on the side of the track. This makes for very complex, inefficient and often dysfunctional operations networks!

Then things started to change. Replacing the 'working block' method of managing train movements, and its associated physical infrastructure of signals and signal boxes, came the European Train Control System (ETCS). This system created a digital 'moving block' by transmitting the actual location of trains via sensors placed close together along the track, in much the same way as aircraft transponders report the position of a plane to air traffic control. Large rail operating centres then used advanced software to calculate the optimum distance between trains, factoring in train speeds, braking distances and stations on the route. The effect was to create a safety buffer zone around each train, while minimising unnecessary distance between trains. When implemented, the ETCS was expected to have a positive effect on safety and general reliability through improved coordination, and also create increases in effective capacity. Conventional railway networks have a mix of high-speed passenger trains and low-speed freight trains, which creates significant problems for the fixed-position 'working block' system. The 'moving blocks' used by the ETCS could increase effective *processing* capacity by around 40 per cent.

The system was rolled out on a number of dedicated high-speed lines, such as the Wuhan–Guangzhou route in China and the TGV route in France. However, the bigger tests were expected to come as moving-block technology was implemented across more complex networks with many more legacy systems to replace. These included projects to install ETCS on trans-European 'corridor' routes connecting different EU countries, specific routes or regions in Australia, Hungary, Italy and New Zealand, and across the entire railway networks of Belgium, Denmark, Germany, Israel and the UK. But even these projects looked small when compared with intended roll-out of ETCS across India's enormous railway network, which moves over 20 million passengers and 3 million tonnes of freight daily. The initial plan was to roll-out the technology across the entire 65,000-km network with a ₹780 billion ($12 billion) contract over six years. However, cost concerns led to scaling back, with the Indian Railway Board instead deciding to undertake a full-scale pilot on a 780-km section of the Delhi–Kolkata route before finalising its future roll-out strategy.

Naufal MQ/Moment/Getty Images

Capacity depends on activity mix

An operation's ability to supply is partly dependent on *what* it is being required to do. For example, a hospital may have a problem in measuring its capacity because the nature of its service may vary significantly. If all its patients require relatively minor treatment with only short stays in hospital, it could treat many people per week. Alternatively, if more of its patients require long periods of observation or recuperation, it could treat far fewer. Output depends on the mix of activities in which the hospital is engaged and, because most hospitals perform many different types of activities, output is difficult (though not impossible!) to predict. Some of the problems caused by variation mix can be partially overcome by using aggregated capacity measures. 'Aggregated' means that different products and services are bundled together in order to get a broad view of demand and capacity. Medium-term capacity management is usually concerned with setting capacity levels in aggregated terms, rather than being concerned with the detail of individual products and services. Although this may mean some degree of approximation, especially if the mix of products or services being produced varies significantly, it is usually acceptable, and is a widely used practice in medium-term capacity management. For example, a hotel might think of demand and capacity in terms of 'room nights per month'. This ignores the number of guests in each room and their individual requirements, but it is a good first approximation.

Capacity depends on the duration over which output is required

Capacity is the output that an operation can deliver *in a defined unit of time*. The level of activity and output that may be achievable over short periods of time is not the same as the capacity that is sustainable on a regular basis. For example, a tax return processing office, during its peak periods at the end (or beginning) of the financial year, may be capable of processing 120,000 applications a week. It does this by extending the working hours of its staff, discouraging its staff from taking vacations during this period, avoiding any potential disruption to its IT systems (e.g. not allowing upgrades during this period), and maybe just by working hard and

Worked example

The impact of activity mix on drone production capacity

A small engineering firm based in Stockholm, Sweden, manufactures a range of three commercial drones – *Vortex*, *Elysia* and *Moln* (cloud). The *Vortex* model can be assembled in 2.5 hours, the *Elysia* in 1.5 hours and the *Moln* in 0.75 hours. The firm has 600 staff hours available for assembly per week. Assuming that the demand mix for the *Vortex*, *Elysia* and *Moln* drones are in a ratio of 2:2:4, the time needed to assemble 2 + 2 + 4 = 8 units is:

$$(2 \times 2.5) + (2 \times 1.5) + (4 \times 0.75) = 11 \textit{ hours}$$

The number of units assembled per week is therefore:

$$\frac{600}{11} \times 8 = 436.4 \textit{ units}$$

If the activity mix changed to a ratio of 3:2:3, the time needed to assemble 3 + 2 + 3 = 8 units is:

$$(3 \times 2.5) + (2 \times 1.5) + (3 \times 0.75) = 12.75 \textit{ hours}$$

Now the number of units assembled per week (i.e. the operation's new capacity) is:

$$\frac{600}{12.75} \times 8 = 376.5 \textit{ units}$$

intensively. However, the capacity that is possible to cope with peak periods is not sustainable over long periods. Staff do need vacations, they cannot work long hours continually, and eventually the information system will have to be upgraded. As such, capacity is taken to be the level of activity or output that can be sustained over an extended period of time.

Capacity depends on the specification of output

Some operations can increase their output by changing the specification of the product or service (although this is more commonly applied to service contexts). For example, during the busy Christmas period, a postal service may effectively reduce its dependability, with next-day delivery rates falling from 95 per cent to 85 per cent. Similarly, accounting firms may avoid long 'relationship-building' meetings with clients during busy periods. The important task is to distinguish between the 'must do' elements of the service that should not be sacrificed and the 'nice to do' parts of the service that can be omitted or delayed in order to increase capacity in the short term.

Capacity 'leakage'

Even after allowing for all the difficulties inherent in measuring capacity, the theoretical capacity of a process (the capacity that it was designed to have) is rarely achieved in practice. Some reasons for this are, to some extent, predictable. Different products or services may have different requirements, so people and machinery will have delays when switching between tasks. Maintenance will need to be performed on machines, while training will be required for employees. Scheduling difficulties could mean further lost time. Not all of these losses are necessarily avoidable; they may occur because of the market and technical demands on the process. However, some of the reduction in capacity can be the result of less predictable events. For example, labour shortages, quality problems, delays in the delivery of bought-in products and services, and machine, or system, breakdown, can all reduce capacity. This loss in capacity is sometimes called 'capacity leakage' and one popular method of assessing this leakage is the overall equipment effectiveness (OEE) measure (see Figure 8.5) that is calculated as follows:

$$\text{OEE} = a \times p \times q$$

where a is the availability of a process, p is the performance or speed of a process, and q is the quality of product or services that the process creates.OEE works on the assumption that some capacity leakage occurs as a result of reduced availability. For example, availability can be lost through time losses such as set-up and changeover losses (when equipment, or people in a service context, are being prepared for the next activity), and breakdown failures (when a machine is being repaired or, in a service context, where employees are being trained, or absent). Some capacity is lost through speed losses such as when equipment is idling (for example when it is temporarily waiting for work from another process) and when equipment is being run below its optimum work rate. In a service context, the same principle can be seen when individuals are not working at an optimum rate – for example, mail-order call-centre employees in the quiet period after the winter holiday season. Finally, not everything processed by an operation will be error free. So, for example, a consulting firm having to re-run some of its analysis caused by data input errors would lose some of its effective capacity.

For processes to operate effectively, they need to achieve high levels of performance against all three dimensions – availability, performance (speed) and quality. Viewed in isolation, these individual metrics are important indicators of performance, but they do not give a complete picture of the process's *overall* effectiveness. And critically, all these losses in the calculation mean that OEE represents the valuable operating time as a percentage of the capacity something was designed to have.

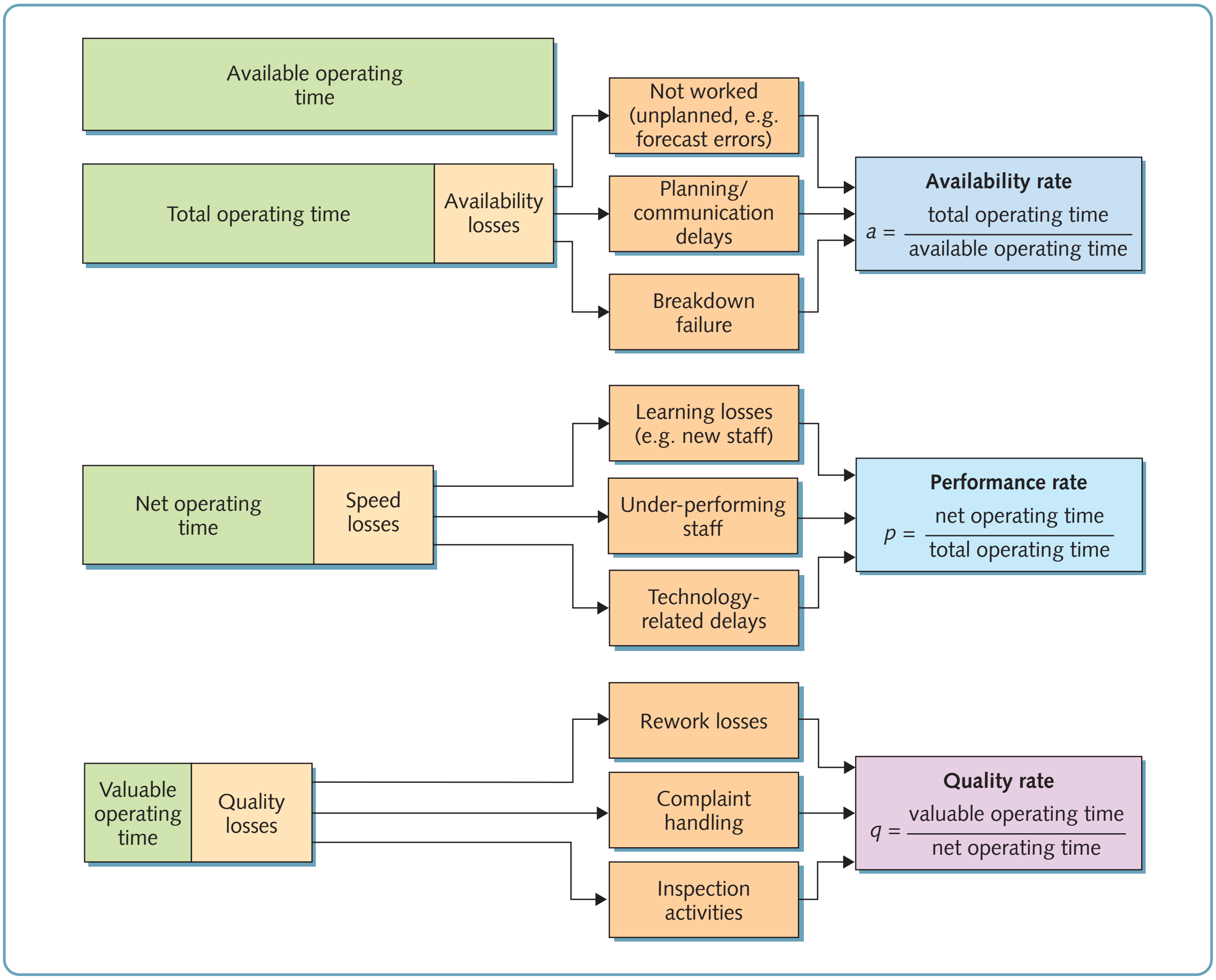


Figure 8.5 Overall equipment effectiveness (OEE)

Case example

Mass transport systems have limited options in coping with demand fluctuation[5]

Anyone who commutes regularly in a large urban area knows the frustration of congestion at peak travelling times. Frustrating for passengers, certainly, but also frustrating for the operations managers who have to cope with sometimes widely variable demand. The mass transport systems on which most large cities depend are faced with one of the most difficult capacity management problems. Their transportation facilities are prone to congestion because of three characteristics of travel: demand varies very significantly over time; supply is relatively fixed over long periods of time; and their service is not storable. The London Underground (usually just called 'the Underground', or 'the Tube' by Londoners) is a good illustration of this. It is the oldest and busiest mass rapid transport (MRT) system in the world. But from the perspective of Transport for London (TfL), who operate the system, an equally pressing issue is how to cope with the fluctuation in demand during each day. In most parts of London Underground, the morning peak time lasts from 07:30 to 09:30 and the evening peak from around 16:40 to 18:30. However, there is some variation, with different parts of the system experiencing higher levels of traffic during different parts of the day. The network is largely closed overnight during the week, with some exceptions for all-night services on Friday and Saturday nights and special occasions.

However, there is some evidence that demand is 'spreading out' as passengers deliberately try to avoid travelling in peak periods. This is a pattern reflected in other mass transport systems. Singapore's Land Transport Authority' (LTA), which runs its system, launched an incentive called Travel Smart Journeys (TSJ), aimed at distributing peak-hour demand more evenly by rewarding commuters along congested areas to consider alternative transport modes or routes. Some commentators see such shifting patterns, particularly the increase in off-peak travel, as reflecting the ways in which people are changing the way they live and work. In London, the shift to more people working from home and the number of self-employed people has increased much faster than the growth in overall employment in the city. To encourage this trend, some MRT systems adopt differential pricing (off-peak lower fares, etc.), such as Singapore's TSR initiative. The other main difficulty for most MRT systems is the more or less fixed capacity of their networks. More and better trains would help, but are expensive. On London Underground there is an ongoing programme of commissioning new trains, some of which allow passengers to move through the train to spread out more efficiently. But there are few options to flex capacity. Essential maintenance is carried out during the nightly closures, and more substantial track maintenance is occasionally done during weekends when demand is lower. Similarly, cleaning work can be performed at night. London Underground has nearly 10,000 night workers. It also helps if passengers can be encouraged to board and exit trains as quickly as possible. The newer trains can help this – they are not much faster between stations, but they do draw passengers off the platform faster, which helps the system to stay fluid at peak times.

Understanding changes in capacity

While many operations are most concerned with dealing with changes in demand, some operations also have to cope with variation in *capacity* (if it is defined as 'the ability to supply'). For example, Figure 8.6 shows the demand and capacity variation of two businesses. The first is a domestic appliance repair service. Here, capacity varies slightly because the field service operatives in the business prefer to take their vacations at particular times of the year. However, it is the demand side of the equation that exhibits more variability, with peak demand approximately twice the level of the low point in demand. The second business is a food manufacturer producing frozen spinach. The demand for this product is relatively constant throughout the year but the capacity of the business varies significantly. During the growing and harvesting season capacity to supply is high, but it falls off almost to zero for part of the year. Yet although the mismatch between demand and capacity is driven primarily by fluctuations in demand in the first case and capacity in the second case, the essence of the capacity management activity is fundamentally similar for both.

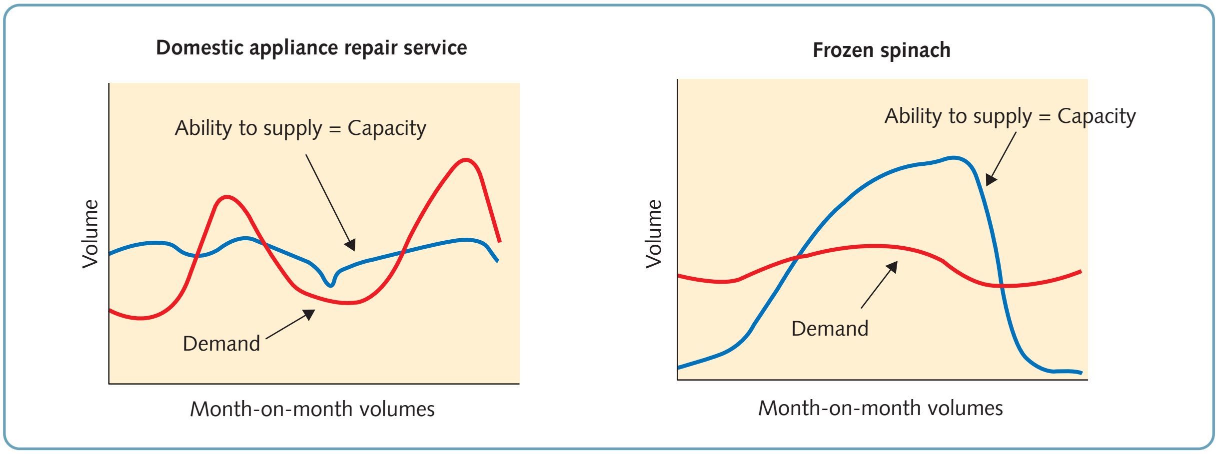


Figure 8.6 Volatility in demand versus volatility in capacity

8.4 Diagnostic question: *Have strategies for managing the 'supply side' been considered?*

> **OPERATIONS PRINCIPLE**
> ***Managing the supply side involves setting base capacity and using 'level' or 'chase' plans to manage the supply of products or services.***

Having determined the capacity of the operation, managers must make a number of important supply-side decisions. These include setting the base capacity level, and then evaluating two key methods of managing supply – *level capacity plans*, where demand fluctuations are ignored and nominal capacity is kept constant, and *chase capacity plans*, where capacity is adjusted to 'chase' fluctuations in demand over time.

Setting base capacity

The most common starting point in managing the supply side is to decide on a 'base level' of capacity and then adjust it periodically up or down to reflect fluctuations in demand. Three factors are important to consider in setting this base level:

1. The operation's performance objectives.
2. Perishability of the operation's outputs.
3. Variability in demand or supply.

The effect of performance objectives on the base level

Base levels of capacity should be set primarily to reflect an operation's performance objectives (see Figure 8.7). For example, setting the base level of capacity high compared to average demand will result in relatively low levels of utilisation and therefore high costs. This is especially true when an operation's fixed costs are high and therefore underutilisation has significant detrimental effects. Conversely, high base levels of capacity result in a capacity 'cushion' for much of the time, so the ability to flex output to give responsive customer service will be

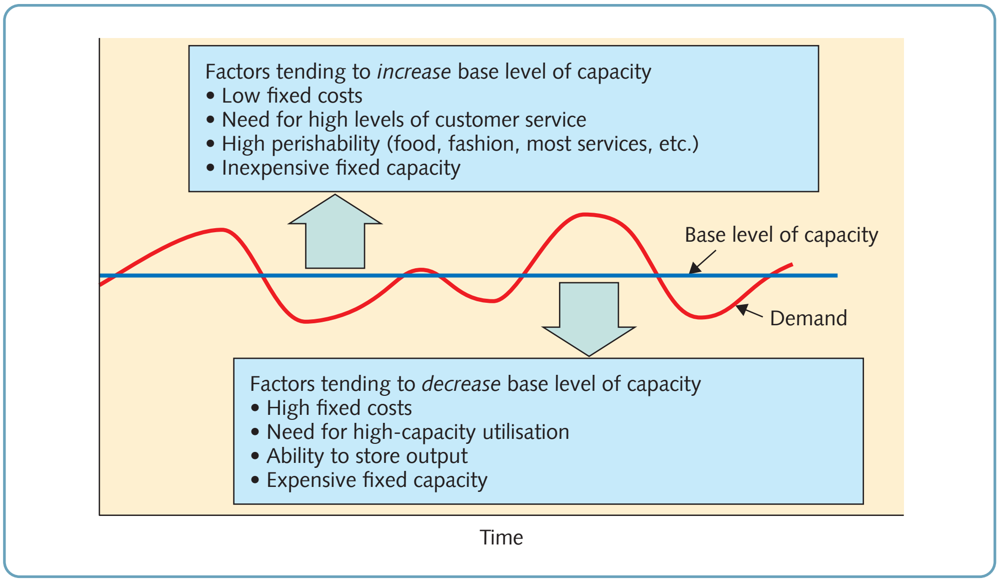


Figure 8.7 The base level of capacity should reflect the relative importance of the operation's performance objectives

enhanced. When the output from the operation is capable of being stored, there may also be a trade-off between fixed capital and working capital in where base capacity level is set. A high level of base capacity can require considerable investment while a lower base level would reduce the need for capital investment but may require inventory to be built up to satisfy future demand, thus increasing working capital. For some operations, building up inventory is either risky because products have a short shelf life (for example perishable food, high-performance computers or fashion items) or because the output cannot be stored at all (for example, most services).

The effect of perishability on the base level

When either supply or demand is perishable, base capacity will need to be set at a relatively high level because inputs to the operation or outputs from the operation cannot be stored for long periods. For example, a factory that produces frozen fruit will need sufficient freezing, packing and storage capacity to cope with the rate at which the fruit crop is being harvested during its harvesting season. Similarly, a hotel cannot store its accommodation services. If an individual hotel room remains unoccupied, the ability to sell for that night has 'perished'. In fact, unless a hotel is fully occupied every single night, its capacity is always going to be higher than the average demand for its services.

The effect of demand or supply variability on the base level

Variability, either in demand or capacity, will reduce the ability of an operation to process its inputs. (The consequences of variability in individual processes were discussed in Chapter 5.) As a reminder, the greater the variability in arrival time or activity time at a process, the more the process will suffer both high throughput times *and* reduced utilisation. This principle holds true for whole operations, and because long throughput times mean that queues will build up in the operation, high variability also affects inventory levels (as illustrated in Figure 8.8). The implication is that the greater the variability, the more extra capacity will be needed to compensate for the reduced utilisation of available capacity.

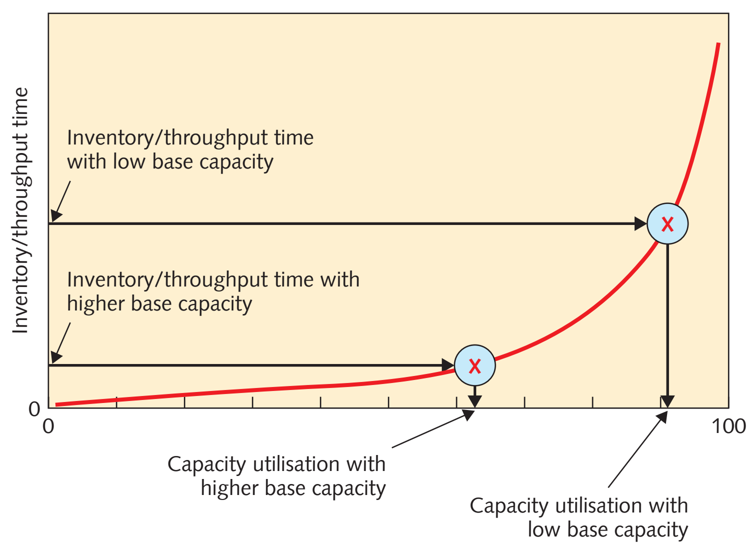


Figure 8.8 The effect of variability on the utilisation of capacity

Level capacity plan

Once base capacity is set, the first alternative supply-side approach is to keep this level fixed throughout the planning period, regardless of the fluctuations in forecast demand. This is called a 'level capacity plan'. It means that the same number of staff or machines operate the same processes and should therefore be capable of producing the same aggregate output in each period. Where non-perishable materials are processed, but not immediately sold, they can be transferred to finished goods inventory in anticipation of later sales. When inventory is not possible, as in most service operations, demand fluctuations are absorbed through underutilisation of the operation's resources and/or failure to meet demand immediately (see Figure 8.9(a)). The more demand fluctuates, the higher the level of inventory or underutilisation there is. Both are expensive but may be considered if the cost of building inventory is low compared with changing output levels; or, in service operations, if the opportunity costs of individual lost sales are very high (for example, in the high-margin retailing of jewellery and in (real) estate agents). Setting capacity below the forecast peak demand level will reduce the degree of underutilisation, but in the periods where demand is expected to exceed capacity, customer service will deteriorate.

Chase (demand) capacity plan

In contrast to level capacity plans, chase capacity plans attempt to match demand patterns closely by varying levels of capacity (as in Figure 8.9(b)). Chase capacity plans are much more difficult to achieve than level capacity plans, as different numbers of staff, different working hours and even different amounts of equipment may be necessary in each period. For this reason, pure chase demand plans are unlikely to appeal to operations producing standard, non-perishable products. Also, where manufacturing operations are particularly capital-intensive, this approach would require a high level of physical capacity, much of which would be used only occasionally. A pure chase plan is more usually adopted by operations that are not able to store their output, such as customer-processing operations or manufacturers of perishable products. It avoids the wasteful provision of excess staff that occurs with a level capacity plan, and yet should satisfy customer demand throughout the planned period. Where output can be stored, the chase demand policy might be adopted in order to minimise or eliminate finished goods inventory, especially if the nature of future demand (in terms of volume or mix) is relatively unpredictable. There are a number of different methods for adjusting capacity (see Table 8.1), although they may not all be feasible for all types of operation.

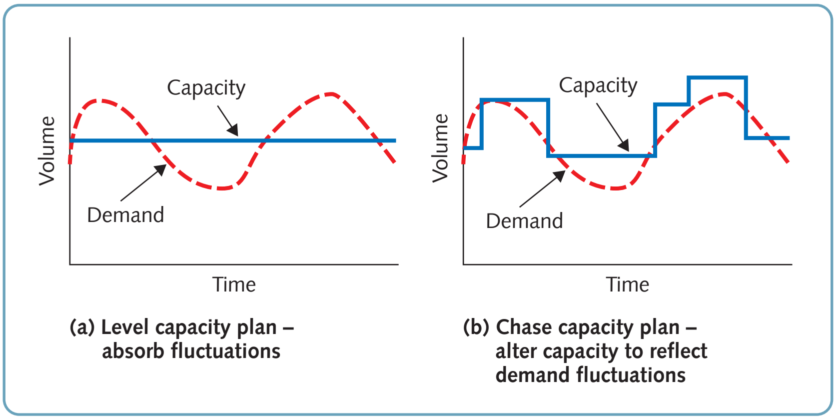


Figure 8.9 **(a) 'Level' capacity plan versus (b) 'Chase' capacity plan**

Table 8.1 Summary of advantages and disadvantages of some methods of adjusting capacity

Method of adjusting capacity	*Advantages*	*Disadvantages*
Overtime – staff working longer than their normal working times	Quickest and most convenient	Extra payment normally necessary and agreement of staff to work; can reduce productivity over long periods
Annualised hours – staff contracted to work a set number of hours per year rather than a set number of hours per week	Without many of the costs associated with overtime the amount of staff time available to an organisation can be varied throughout the year to reflect demand	When very large and unexpected fluctuations in demand are possible, all the negotiated annual working time flexibility can be used up before the end of the year
Staff scheduling – arranging working times (start and finish times) to vary the aggregate number of staff available for working at any time	Staffing levels can be adjusted to meet demand without changing job responsibilities or hiring in new staff	Providing start and finish (shift) times that both satisfy staff's need for reasonable working times and shift patterns as well as providing appropriate capacity can be difficult
Varying the size of the workforce – hiring extra staff during periods of high demand and laying them off as demand falls	Reduces basic labour costs quickly	Hiring costs and possible low productivity while new staff go through the learning curve; lay-offs may result in severance payments and possible loss of morale in the operation and loss of goodwill in the local labour market
Using part-time staff – recruiting staff who work for less than the normal working day (at the busiest periods)	Good method of adjusting capacity to meet predictable short-term demand fluctuations	Expensive if the fixed costs of employment for each employee (irrespective of how long he or she works) are high
Skills flexibility – designing flexibility in job design and job demarcation so that staff can transfer across from less-busy parts of the operation	Fast method of reacting to short-term demand fluctuations	Investment in skills training needed and may cause some internal disruption
Subcontracting/ outsourcing – buying, renting or sharing capacity or output from other operations	No disruption to the operation	Can be very expensive because of subcontractor's margin and subcontractor may not be as motivated to give the same service, or quality; also, a risk of leakage of knowledge
Changing output rate – expecting staff (and equipment) to work faster than normal	No need to provide extra resources	Can only be used as a temporary measure, and even then can cause staff dissatisfaction, a reduction in the quality of work, or both

Zero-hours contracts

A zero-hours contract is one where an employer does not offer any guarantee of a specific number of hours of work for a worker. Neither is any person working under this type of contract under any obligation to accept those hours when they are offered. It can therefore be seen as a method of varying effective capacity for 'chase demand' approaches. The prevalence of zero-hours arrangements, exactly how they are defined and their legal status varies significantly between countries. For example, at the time of writing they are allowed in Hong Kong, Malaysia, Norway, Singapore, UK and the US (though not typically referred to as zero-hours contracts), allowed but heavily regulated in the Netherlands and Sweden, and not generally allowed in Austria, Belgium, China, Czech Republic, France, Germany, Italy and Spain. Zero-hours contracts are controversial because under such schemes people may have no idea about how many hours they will be working (and therefore the wages that they will be earning). But the flexibility may be valued by some people, notwithstanding the fluctuating incomes.

8.5 Diagnostic question: *Have the consequences of capacity management decisions been considered?*

> OPERATIONS PRINCIPLE
> ***Most organisations mix demand-side (demand management and yield management) and supply-side (level and chase plans) capacity management strategies to maximise performance.***

When making capacity management decisions, managers are attempting to balance the need to provide a responsive and customer-oriented service with the need to minimise costs. For this reason, most organisations choose to follow a mixture of the approaches outlined in this chapter. For example, an accounting firm may seek to bring forward some of its peak demand by offering discounts to selected clients (demand management plan). Capacity may also be increased through the use of outsourced suppliers during the busiest months of the year (chase capacity plan). However, some capacity may be constrained (for example, specialist advisory services offered by the firm) and therefore clients may still experience delays during high-demand periods (level capacity plan).

Before an operation adopts one or more of the three 'pure' capacity management plans (demand management, level capacity or chase capacity), it should examine the likely consequences of its decisions. Three methods are particularly useful in this assessment:

- Factoring in predictable versus unpredictable demand variation.
- Using cumulative representations of demand and capacity.
- Taking a longitudinal perspective that considers short- and long-term outlooks.

Factoring in predictable versus unpredictable demand variation to plan capacity

Of critical importance in deciding on the most appropriate mix of capacity management strategies is an understanding of the balance between predictable and unpredictable variation in demand. When demand is stable and predictable, the life of an operations manager is relatively easy! If demand is changeable, but this change is predictable, capacity adjustments may be needed but at least they can be planned in advance. With unpredictable variation in demand, if an operation is to react to it at all, it must do so quickly; otherwise the change in capacity will have little effect on the operation's ability to deliver products and service as needed by its customers. Figure 8.10 illustrates how the objectives and tasks of capacity management vary, depending on the balance between predictable and unpredictable variation.

> OPERATIONS PRINCIPLE
> ***Capacity management decisions should reflect both predictable and unpredictable variations in demand.***

Using cumulative representations to plan capacity

When an operation's output can be stored (in contrast to yield management, where it cannot), a useful method of assessing the feasibility and consequences of adopting alternative capacity

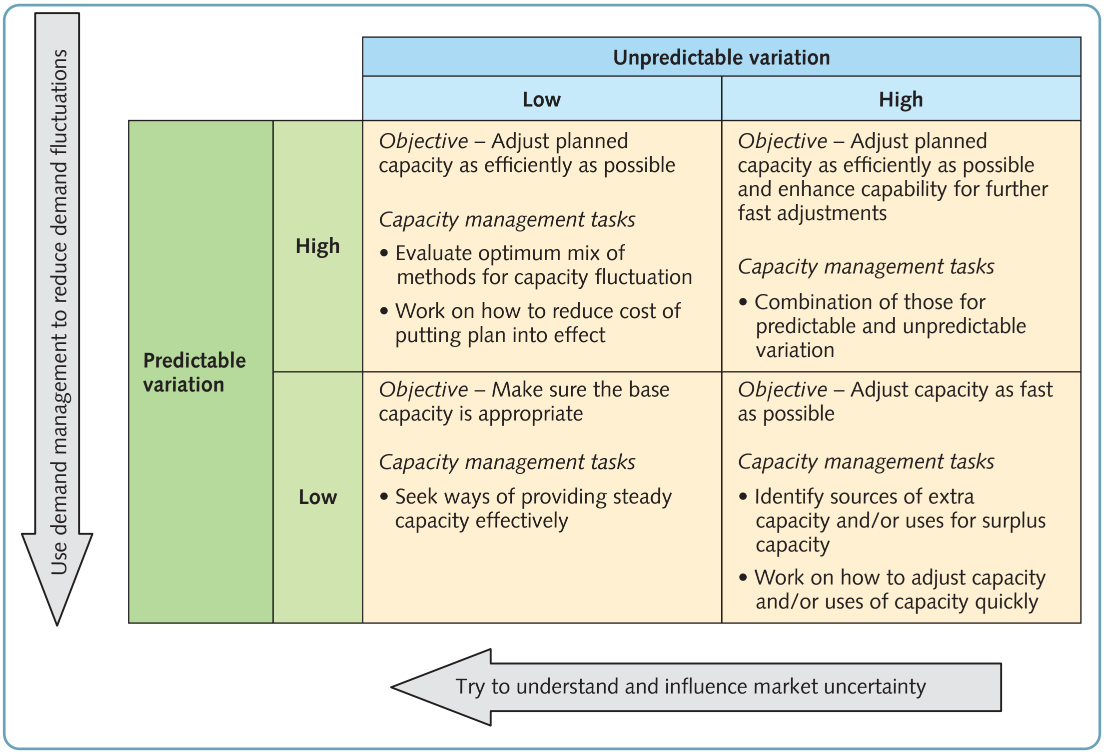
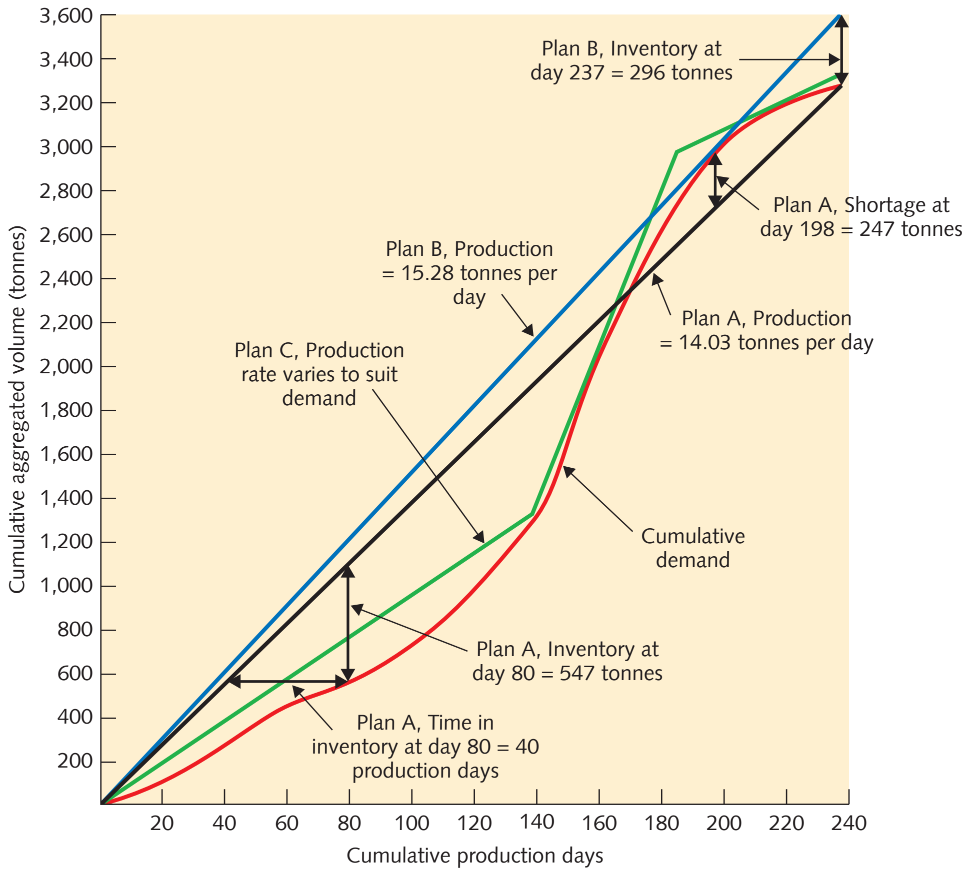


Figure 8.10 Capacity management objectives and tasks depend on the mixture of predictable and unpredictable variation in demand

plans is to plot both the cumulative demand on an operation and its cumulative ability to supply, over time. For example, Figure 8.11 shows the forecast aggregated demand for a chocolate factory. Demand for its products in retailers is greatest in December. To meet this demand and allow time for the products to work their way through their supply chain, the factory must supply a demand that peaks in September. But the cumulative representation of demand against available supply time (productive days) shown in Figure 8.11 reveals that, although total demand peaks in September, because of the restricted number of available productive days, the peak demand per productive day occurs a month earlier in August. It also shows that the effective fluctuation in demand over the year is even greater than it seems. The ratio of monthly peak demand to monthly lowest demand is 6.5:1, but the ratio of peak to lowest demand per productive day is 10:1. Demand per productive day is more relevant to operations managers, because productive days represent the 'ability to supply'.

The feasibility and consequences of a capacity plan can be assessed on this basis. Figure 8.11 also shows a level capacity plan (A) that assumes production at a rate of 14.03 tonnes per productive day. This meets cumulative demand by the end of the year, and there is theoretically no over-capacity or under-capacity. However, if one of the aims of the plan is to supply demand when it occurs, the plan is inadequate. Up to around day 168, the line representing cumulative production is above that representing cumulative demand. This means that at any time during this period, more product has been produced by the factory than has been demanded from it. In fact, the vertical distance between the two lines is the level of inventory at that point in time. So, by day 80, 1,122 tonnes have been produced but only 575 tonnes have been demanded. The surplus of production above demand, or inventory, is therefore 547 tonnes. When the cumulative demand line lies above the cumulative production line, the reverse is true. The vertical distance between the two lines now indicates the shortage, or lack of supply. For example, by day 198, 3,025 tonnes have been demanded but only 2,778 tonnes produced. The shortage is therefore 247 tonnes.



	J	F	M	A	M	J	J	A	S	O	N	D
Demand (tonnes/month)	100	150	175	150	200	300	350	500	650	450	200	100
Production days	20	18	21	21	22	22	21	10	21	22	21	18
Demand (tonnes/day)	5	8.33	8.33	7.14	9.52	13.64	16.67	50	30.95	20.46	9.52	5.56
Cumulative days	20	38	59	80	102	124	145	155	176	198	219	237
Cumulative demand	100	250	425	575	775	1075	1425	1925	2575	3025	3225	3325

Figure 8.11 Cumulative representation of demand and three capacity plans

> **OPERATIONS PRINCIPLE**
> ***For any capacity plan to meet demand as it occurs, its cumulative production line must always lie above its cumulative demand line.***

For any capacity plan to meet demand as it occurs, its cumulative production line must always lie above the cumulative demand line. This makes it a straightforward task to judge the adequacy of a plan, simply by looking at its cumulative representation. An impression of the inventory implications can also be gained from a cumulative representation by judging the area between the cumulative production and demand curves. This represents the amount of inventory carried over the period. In Figure 8.11, level capacity plan (B) is feasible because it always ensures enough production to meet demand at any time throughout the year. However, inventory levels are high using this plan. It may even mean that the chocolate spends so much time in the factory's inventory that it has insufficient shelf life when it arrives at the company's retail customers (the more perishable the product, the greater the risk of this type of plan). Assuming a 'first-in-first-out' inventory management principle, the time a product stays in inventory will be represented by the horizontal line between the demand at the time it is 'demanded' and the time it was produced.

The gap between demand and supply over time can be reduced by adopting a chase (demand) capacity plan, shown as (C) in Figure 8.11. This reduces the average time products spend as inventory, but incurs costs associated with changing capacity levels. Usually, the marginal cost of making a capacity change increases with the size of the change. For example, if the chocolate manufacturer wishes to increase capacity by 5 per cent, this can be achieved by requesting its staff to work overtime – a simple, fast and relatively inexpensive option. If the change is 15 per cent, overtime cannot provide sufficient extra capacity and temporary staff will need to be employed – a more expensive solution, which also would take more time. Increases in capacity of above 15 per cent might only be achieved by subcontracting some work out. This is likely to be even more expensive.

Taking a longitudinal perspective to plan capacity

Although planning capacity levels in advance, and even planning how to respond to unexpected changes in demand, is an important part of capacity management, it does not fully reflect the dynamic nature of the activity. Capacity management must react to *actual* demand and *actual* capacity as it occurs. Period by period, managers consider demand forecasts, understanding of current capacity and (if outputs can be stocked) how much inventory or queues of customers in the process have been carried forward from the previous period. Based on all this information, they make plans for the following period's capacity. During the next period, demand might or might not be as forecast and the actual capacity of the operation might or might not turn out as planned (because of the capacity leakage discussed earlier). But whatever the actual conditions during that period, at the beginning of the next period the same types of decisions must be made, in the light of the new circumstances. Figure 8.12 shows how this works in practice.

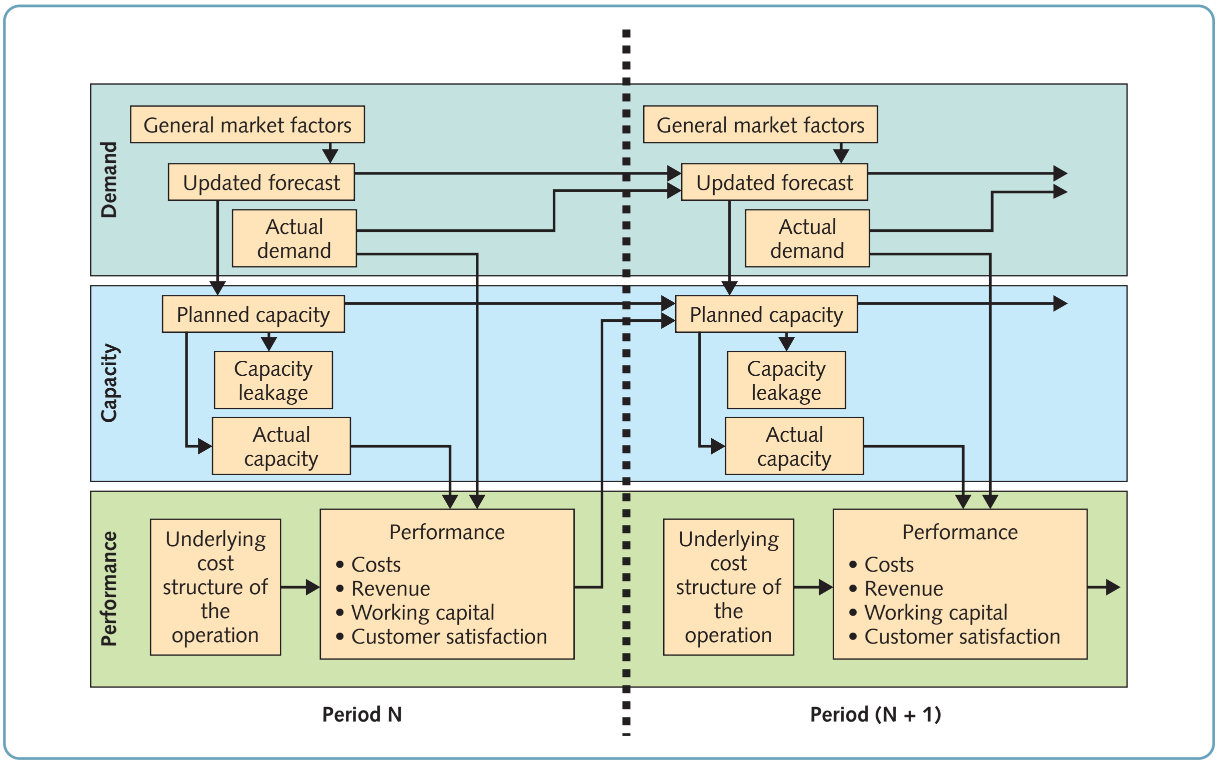


Figure 8.12 How capacity should be controlled – the dynamics of capacity management

The success of capacity management is generally measured by some combination of costs, revenue, working capital and customer satisfaction (which goes on to influence revenue). This is influenced by the actual capacity available to the operation in any period and the demand for that period. If capacity is in excess of demand, customer demands can be met, but underutilised capacity, and possibly inventory, will increase costs. If capacity is less than demand, the operation's resources will be fully utilised, but at the expense of being unable to meet all demand. The effect is that customers will be waiting for products and service (or in many cases will be switching to competitors who are able to provide what they need!). However, some operations are better able to cope than others with any mismatch between actual capacity and actual demand. If the underlying cost structure of the operation is such that fluctuations in output level have relatively little effect on costs, then the operation will be less sensitive to errors in capacity management. However, overriding other considerations already discussed, the capacity strategy adopted is often influenced by any differences between the long- and short-term outlooks for demand (see Figure 8.13).

Case example

United drags passenger off its plane[6]

When footage shot by a fellow passenger showed a bloodied and unconscious man being pulled off of a flight at O'Hare International Airport in Chicago, the clip caused a sensation on social media and later on mainstream news outlets. It also called down a barrage of criticism on United Airlines, the offending carrier. The incident began when United overbooked the flight (a problem made worse because at the last minute it decided to fly four members of staff to a connection point, so they could staff another flight). The airline decided that it needed to bump four passengers to make way for them. This is common practice; when a flight is overbooked, the first step is to offer a financial encouragement for passengers to take a later flight. In this case, passengers were first offered $400, overnight hotel accommodation and a flight the following day. No one accepted and the offer was increased to $800. Still no one accepted the offer, so a manager announced that passengers would be selected to leave the flight, with frequent fliers and business-class passengers being given priority. The first two people selected agreed to leave the plane. The third person selected (as it happened, the wife of the man who was later dragged off forcibly) also agreed. However, when the fourth person was approached, he refused, saying that he was a doctor and had to see patients in the morning. Eyewitnesses said the man was 'very upset' and tried to call his lawyer. So, instead of selecting another passenger, or increasing its offer (it could have offered a maximum of $1,350), security staff were called. The encounter with the security staff concluded with the man being wrenched from his seat onto the floor, after which he was hauled down the aisle, blood covering his face.

It is not uncommon for airlines to sell more tickets than they have seats, under the assumption that some passengers will either fail to show or cancel at the last minute. It is claimed that, because of fierce competition and to avoid flying with empty seats, so-called 'denied boarding' incidents are becoming more common. As the incident gained publicity, the CEO of United Airlines, Oscar Munoz, said that employees '*followed established procedures*', but he was '*upset to see and hear about what happened*', although the passenger had been '*disruptive and belligerent*'. When the passenger refused to leave the plane voluntarily, he said, the staff were '*left with no choice but to call security officers to assist in removing the customer from the flight*'. Travel expert, Simon Calder, said that the airline was technically within its rights: '*The captain is in charge of the aircraft. And if he or she decides that someone needs to be offloaded, that command has to be obeyed. From the moment that the unfortunate individual in this case said, "I'm staying put", he became a disruptive passenger. Officials were legally entitled to remove him, and as the videos show, he was dragged from the plane. It appears from the evidence that the law was broken – by him, not by the airline. But I would be surprised if United pressed charges.*'

Long-term outlook for volume \ Short-term outlook for volume	Decreasing below current capacity	Level with current capacity	Increasing above current capacity
Decreasing below current capacity	Reduce capacity (semi) permanently: for example, reduce staffing levels; reduce supply agreements	Plan to reduce capacity (semi) permanently: for example, freeze recruitment; modify supply agreements	Increase capacity temporarily: for example, increase working hours and/or hire temporary staff; modify supply agreements
Level with current capacity	Reduce capacity temporarily: for example, reduce staff working hours; modify supply agreements	Maintain capacity at current level	Increase capacity temporarily: for example, increase working hours and/or hire temporary staff; modify supply agreements
Increasing above current capacity	Reduce capacity temporarily: for example, reduce staff working hours, but plan to recruit; modify supply agreements	Plan to increase capacity above current level; plan to increase supply agreements	Increase capacity (semi) permanently: for example, hire staff; increase supply agreements

Figure 8.13 Capacity management strategies are partly dependent on the long- and short-term outlooks for volume

Critical commentary

- A key source of debate in capacity management concerns the relative benefits of superior forecasting versus superior operational flexibility. On one side are those who argue that effective capacity planning is impossible without high-quality forecasts. While patterns of demand may sometimes be hard to decipher, a combination of qualitative and quantitative methods, drawing on a range of data sources, will always give useful insights. Those organisations that do not invest resources in attempting to better predict demand for their products or services are 'flying blind'. The counterargument is that demand is inherently uncertain, at least for many organisations, and investing significantly in forecasting is often a waste of time and money. Those taking this position argue instead for investments in operational flexibility that enable demand to be met, irrespective of what it is. Furthermore, some see the whole idea of forecasts as fundamentally risky because they give a false sense of security about what will happen in the future, when of course they can only tell us what *might* happen. These arguments both have merit, but in practice most operations need to find a balance between improving forecasts and coping with uncertainty. Trying to get forecasts right (or better) has particular value where the operation finds it difficult or impossible to react to unexpected demand fluctuations in the short term. On the other hand, organisations with intrinsically uncertain markets may focus more on developing fast and flexible processes to compensate for the difficulty in obtaining accurate forecasts. Similarly, when the cost of not meeting demand is very high, operations must typically rely on their responsiveness rather than on the accuracy of any forecasts.
- For such an important topic, there is surprisingly little standardisation in how capacity is measured. Not only is a reasonably accurate measure of capacity needed for operations management, it is also needed

to decide whether it is worth investing in extra physical capacity. Yet not all practitioners would agree with the way in which capacity has been defined or measured in this chapter (although it does represent orthodox practice). One school of thought is that whatever capacity efficiency measures are used, they should be useful as diagnostic measures that can highlight the root causes of inefficient use of capacity. The idea of overall equipment effectiveness (OEE) described earlier is often put forward as a useful way of measuring capacity efficiencies.

- Another point of controversy in capacity management concerns the use of varying staff levels. To many, the idea of fluctuating the workforce to match demand, either by using part-time staff or using zero-hours contracts, or by hiring and firing, is more than just controversial: it is regarded as unethical. It is the responsibility of businesses, they argue, to engage in a set of activities that are capable of sustaining employment at a steady level. In addition, hiring people on a short-term contract, in practice, leads to them being offered poorer conditions of service and leads to a state of permanent anxiety as to whether they will keep their jobs. On a more practical note, it is pointed out that, in an increasingly global business world where companies often have sites in different countries, those countries that allow hiring and firing are more exposed to 'downsizing' than those where legislation makes this difficult. In addition, knowledge-oriented organisations face significant risk when using contracts that do not tie in workers in a meaningful way. In other words, talent leaves at the end of the day and a firm can't be sure it will return in the morning!

SUMMARY CHECKLIST

- ☐ Is the importance of effective capacity management fully understood?
- ☐ Are capacity management decisions coherent across different timescales?
- ☐ Are demand patterns fully understood over different time frames?
- ☐ What potential is there for making unpredictable variability more predictable through better forecasting?
- ☐ Has a combination of qualitative and quantitative methods been considered to maximise the quality of demand forecasts?
- ☐ Is it desirable to change the current demand patterns? If so, what methods have been considered to achieve this?
- ☐ Is the operation's current capacity measured?
- ☐ If so, are all the assumptions inherent in the measurement of capacity made fully explicit?
- ☐ What capacity 'leakage' is normal, and have options for minimising capacity leakage been explored?
- ☐ Is there scope for using the overall equipment effectiveness (OEE) measure of capacity?
- ☐ Has the operation's base level of capacity been determined and have all factors influencing this level been considered?
- ☐ Have methods of managing the 'supply side' using a level capacity plan, a chase (demand) capacity plan, or a combination of the two, been fully explored and assessed?
- ☐ What is the balance between predictable variation and unpredictable variation in demand and capacity?
- ☐ Is there scope for using cumulative representations of demand and capacity for planning purposes?
- ☐ Does the operation take a longitudinal perspective in its capacity management by examining short-term and longer-term outlooks?
- ☐ Is the method of deciding period-by-period capacity levels effective? And how does it reflect previous experience?

Case study

FreshLunch

(This case was co-authored with Vaggelis Giannikas at School of Management, University of Bath)

'Carlos, are you ready to head out then?', Antônia called across the office. *'Too right! After the morning I've had, I could do with the break!'* Carlos laughed, as he grabbed his wallet and sunglasses. As the two headed towards the lift (elevator), they entered into a deep conversation. *'So, what do you fancy Antônia?'* After a short pause to think, Antônia responded, *'Well, we could go to Byōōdē – their red curry is definitely one of the best around and the Pad Thai's pretty good too. Or Pollo Picante? I had the chicken with chimichurri sauce the other day and it was really good.'* As the lift descended from the 32nd floor of their building, the conversation continued. *'If we're looking for something hot Antônia, I guess we could also try the new Indian thali place? Rebecca went there last week and said it was excellent, though she did mention it was pretty slow service. Besides, it's such a nice day, maybe we could have something cold instead? There's the sushi at Kazoku – it's so fresh and there's loads of choice.'* Antônia thought for moment. *'Well I'm absolutely fine not having anything hot, but I'm not in the mood for sushi today to be honest.'* There were just so many options, she reflected. *'How's about FreshLunch instead?'* Carlos smiled as they walked out into bright May sunshine. *'Sounds like a good plan to me!'*

A few blocks away, Sofia had already been serving customers at FreshLunch for an hour and, as usual, things were picking up quickly towards the lunchtime rush. She was a chemist by training and had spent the first six years after her graduation working for a large multinational in a research laboratory based in Norway. But her passion was food and fresh produce. Having completed a part-time executive MBA, Sofia had changed her career direction dramatically and set up as a restaurant owner. Knowing how demanding busy customers were, Sofia had established FreshLunch utilising the techniques she learned from her studies in an attempt to manage her operation effectively. FreshLunch had taken the traditional cafeteria-style approach often found in universities and large hotels, and developed a process that offered quality, variety and speed. The process was simple, involving five sequential steps from order placement to delivery (see Figure 8.14). Having collected any items from the fridge (drinks and sweet treats), the customer first had to choose the base for their meal – a selection between rice and couscous. Next, they would choose their main protein, including chicken, lamb, steak, salmon or grilled vegetables. Then, two sides were selected to accompany the main meal, from a choice of around ten different plates full of vegetables and salads. Finally, dressings and sauces were available before the customer moved to payment at the end of the process. For each step, the customer moved along with their tray, which was passed from one 'assembler' to the next until it was put in a bag and handed to the customer by the cashier.

Since its opening, FreshLunch had always been very busy around lunchtime, with long queues created at the counter, some of which extended outside the restaurant itself. Sofia was happy that her hard work over the past three years was paying off. However, she was beginning to appreciate that these long queues were not translating into the profits needed to create a sustainable business. Sofia tried to think again about how some of the things she had learned during her MBA programme might help her tackle the situation. But while she had applied so much to the business and made plenty of improvements, she felt increasingly like she was too close to the problems to see them clearly. *'What I need'*, she reflected, *'is a fresh pair of eyes'*. That evening, scrolling through her social media, Sofia noticed that her friend Zuri had just posted an interesting piece on the challenges of demand forecasting during 'X-factor' events. *'Now why didn't I think of her before?'*, she thought. After their time together on the MBA programme, Zuri's small consulting business, which helped its customers analyse and improve their operations, had grown substantially, so she was clearly doing something right. Sofia dialled her number and after a few rings, Zuri answered: *'Sofia! Long time, no chat! How are you and how's things in the restaurant business?'*

Fifteen minutes later, their conversation had turned from generally catching up to FreshLunch. *'I find it really hard to predict what my customers are going to choose every day and I often end up having to throw away quite a lot of food. I tried cooking fewer portions but then I had a lot of annoyed customers and I can't risk bad reviews!'* Zuri already had a few ideas in mind but decided to ask a few more questions to understand the business better: *'What about the customers? Do you know if they are happy with*

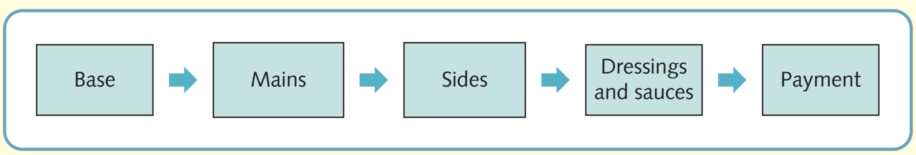


Figure 8.14 FreshLunch process

★★★★★ 4 hours ago

Lovely salads and very friendly staff. We really enjoyed our meal and it's good if you have a mix of meat eaters and vegetarians. We'll be back!

★★★ 6 hours ago

The food is really tasty as long as you don't mind waiting. If you're in a hurry, go somewhere else. FreshLunch need to sort out their queues.

★★★★ 1 day ago

Love the relaxed atmosphere and you never feel rushed making your choices. The food is always really fresh (so the name of the place is spot-on!) and we like how simple the process is.

★★★★★ 4 days ago

I love this place. The food is perfect and simple, and the staff are on it. I try to go about 11.45 so I don't get stuck in the long queues.

★★★ 1 week ago

Pretty good food and reasonable price. Wish they had a bigger selection of mains and sides as I go regularly and am getting a bit bored with the same stuff.

★★★★ 1 week ago

The food is great and I love the roasted cauliflower salad. A bit inconsistent on the steaks and lamb – seems like they get cooked differently every time I order them.

★ 2 weeks ago

Never again! We waited for ages and when we were finally served, they didn't have any chicken and half of the side options weren't available either. Loads of other better options nearby!

★★★★ 3 weeks ago

What a great little find. I had the salmon with baby spinach and a beetroot and feta salad. My colleague had chicken with some lentils and a Greek salad. We both really liked everything. Fast becoming one of our go-to places. Shame they don't serve breakfast, as we're often in work for 6am.

★★★★ 1 month ago

Best place for fresh salads in the area. The slow service isn't ideal, but the food's worth waiting for.

Figure 8.15 Recent customer reviews of FreshLunch

what they get?', she asked. *'I love how busy we are'*, replied Sofia, *'but some customers have already started posting negative reviews due to the long waits. To be fair, it's normally only 20 minutes but since most people only have an hour for lunch, I completely understand where they're coming from.'* Zuri quickly did a search for FreshLunch and began scrolling through some of the most recent customer reviews (see Figure 8.15).

As Zuri read, Sofia continued: *'Unfortunately, the huge rent is kind of killing me to be honest. You can imagine how expensive it is to rent even a small place in the centre of London. It's also hard to find good cooks and waiters, as FreshLunch can only offer them contracts for 50–80 per cent of their time.'* Zuri leaned back on her chair. *'OK Sofia, let me have a think about this over the weekend and I'll get back to you with my thoughts. It's been great catching up.'* As she took a sip of her drink, she thought about how this could be a good exercise for the new associates in her company.

On Monday morning, it was Zuri's turn to pick up the phone and call her friend. *'Sofia, I'd really like to help you with this. We've just hired a small group of young associates and I'd be happy to assign them to work with you. From my perspective, it would give me a chance to see these guys in action and get a feel of how they work as a team, before I set them off on the paid jobs. And for you, it'd be some free consulting – feels like a win–win, right?! I can give them a little bit of supervision, but not much as it's basically pro bono (without charge).'* Sofia was delighted: *'Wow Zuri, that would be fantastic! And I'm happy to give them a bit of advice on any client interaction issues that come up during this.'* Zuri knew that could prove extremely helpful. Sofia had always been excellent at giving constructive feedback. *'Good thinking – that would be brilliant. To get this started, can you please send me over some information about FreshLunch? I will ask my associates to get in touch with a data spec first thing tomorrow.'*

Over the rest of the week, Sofia collected the information that Zuri's team had requested for the project. She started with some basic information, such as opening times (11.00–15.30) and the daily menu (see Figure 8.16).

FreshLunch

✦ OPEN EVERY WEEKDAY 11.00 – 15.30 ✦

STEP 1: Choose a base

COUSCOUS
RICE

STEP 2: Choose a main

CHINESE CHICKEN	£7.95
FLANK STEAK	£8.95
MEDITERRANEAN CHICKEN	£7.95
ORGANIC GRILLED VEGETABLES	£7.45
PERSIAN LAMB	£8.95
THAI SALAMON	£8.95

Amritansh Sagar/Shutterstock

STEP 3: Choose two sides

BABY SPINACH
BEETROOT AND FETA SALAD
CHEESE CROQUETTES
CHICKPEAS WITH SPICY DRESSING
GNOCCHI NAPOLITANA
GREEK SALAD
LENTILS, HALLOUMI AND HERBS
OVEN BAKED SWEET POTATO FRIES
PEA AND MINT SALAD
ROASTED CAULIFLOWER AND BROCCOLI SALAD

STEP 4: Dressings and sauces

GREEK VINAIGRETTE
HONEY MUSTARD
LEMON DRESSING
NONNA'S PESTO
PEPPERCORN SAUCE
WILD GARLIC MAYO
YOGHURT AND MINT

✦ Sweet desserts ✦

CARROT CAKE	£3.45
GIANT MACARON (ALL FLAVOURS)	£3.45
HAZELNUT CHOCOLATE MOUSSE	£3.45
SUMER FRUIT SALAD	£3.45

✦ Beverages ✦

ELDERFLOWER PRESSÉ	£2.95
FRESH APPLE JUICE	£3.95
FRESH ORANGE JUICE	£3.95
GINGER BLITZ SMOOTHIE	£5.45
ICELANDIC STILL WATER	£1.85
MANGO AND CARROT SMOOTHIE	£5.45
RHUBARB PRESSÉ	£2.95
SUMMER FRUIT SMOOTHIE	£5.45

Dishes may contain allergens. If you have any dietary requirements, please speak to a member of staff.

Figure 8.16 FreshLunch daily menu

Note: some sides change periodically, based on seasonality and popularity

Table 8.2 FreshLunch demand across a 'typical' day

Time slot	*Meals*	*Customers (measured in number of receipts)*
11.00–11.30	10	10
11.30–12.00	20	19
12.00–12.30	38	30
12.30–13.00	89	68
13.00–13.30	154	121
13.30–14.00	92	66
14.00–14.30	24	22
14.30–15.00	12	11
15.00–15.30	4	4

She also spent some time putting together information that could be used to analyse the demand patterns for FreshLunch. Luckily, she had recently installed a software package that allowed her to collect and analyse point-of-sale (POS) data. Now Sofia felt she was actually beginning to make some use of it. She remembered one of her professors talking about organisations 'drowning in data' and was starting to appreciate what she meant! To keep it simple, she began with what felt to her like a typical day and broke it down into 30-minute time slots (see Table 8.2). She included information on how many customers typically visited, but also the number of actual meals prepared, as some customers would order more than one meal. In addition, Sofia was asked by the associates to provide some information about daily sales of meals over recent weeks (See Table 8.3).

At Zuri's firm, Bankole Consulting, the associates were looking forward to working on the project. For the group, it was a chance to get stuck into the world of consulting and prove that they were ready to step up to working with the firm's paying clients. It was also an opportunity to repay some of that faith that Zuri had shown them when making her hires. After an initial meeting with Zuri, they started analysing the information that Sofia had sent over. They also decided to pay a visit to FreshLunch to get some first-hand experience of the operation. With Sofia's approval, they behaved as normal customers, queuing for food, ordering and then eating at the bench by the window. During their visit, they drew the layout of the main floor of the restaurant (see Figure 8.17), as this could be useful for their discussions with Zuri. They also looked at the basement area – used for storing ingredients, crockery and utensils – and an area on the upper floor, which functioned as a break room and office space for the shift manager.

During their visit, the team collected data on the size of the queues – something not currently captured by Sofia. There were two parts of the queue. The *assembly queue* was formed between the point a customer started giving their order and the point of payment at the cashier. This was always moving relatively quickly as the employees were used to receiving orders and serving the customers efficiently. The second and more concerning part of the queue was the one formed by people *waiting* to place their order. The team noted the size of the queue at 15-minute intervals over a 3-hour period

Table 8.3 FreshLunch meals sold over a three-week period

Day	*Med chicken*	*Chinese chicken*	*Flank steak*	*Persian lamb*	*Thai salmon*	*Grilled veg*	*TOTAL*
Monday	139	44	28	83	105	155	554
Tuesday	83	33	34	66	57	57	330
Wednesday	102	53	44	89	75	80	443
Thursday	80	30	33	64	60	63	330
Friday	133	55	83	139	100	46	556
Monday	134	62	29	84	95	157	561
Tuesday	84	40	30	67	48	63	332
Wednesday	121	36	44	89	76	81	447
Thursday	85	34	34	68	57	60	338
Friday	129	62	84	138	101	45	559
Monday	141	56	30	85	96	158	566
Tuesday	88	39	34	68	51	61	341
Wednesday	104	55	44	90	77	81	451
Thursday	78	36	34	70	61	61	340
Friday	136	56	87	136	102	47	564
							6,712

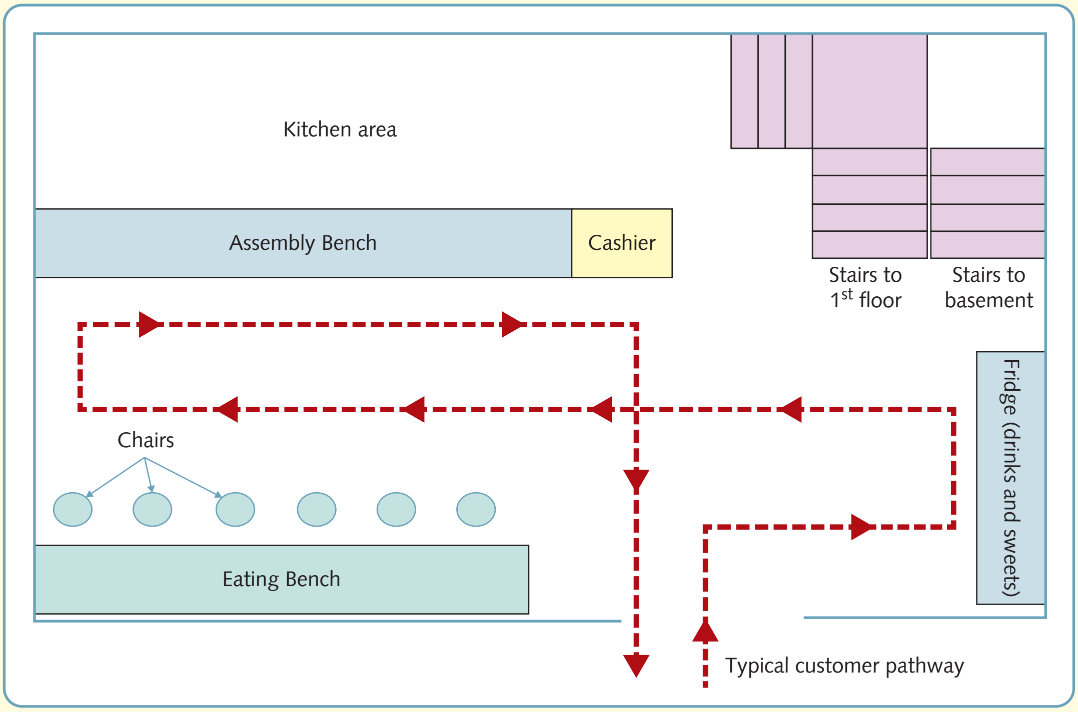


Figure 8.17 FreshLunch restaurant layout

Table 8.4 Queue size during associates' visit to FreshLunch

Time	11.30	11.45	12.00	12.15	12.30	12.45	13.00	13.15	13.30	13.45	14.00	14.15	14.30
Queue	0	0	1	3	6	11	15	21	18	8	3	2	0

(see Table 8.4). Most of the people in the queue spent their time talking to each other, checking their phones, or looking at the printed menus to decide what to order. Zuri's associates also noticed that some people left the queue (nine in total during the three busiest periods) and others decided not to join the queue (around 4–5 in the 15-minute intervals when the queue exceeded 10 customers). Finally, the associates talked to Sofia and her employees to gain further insights into FreshLunch (see following quotes).

Selected quotes from staff at FreshLunch

'This 90 minutes during lunch is hell. It is always busy and there is nothing worse than hungry people.'

'The space upstairs is quite nice, but we hardly ever get the time to use it. By the time things calm down, our shift is over.'

'It's nice to get a free meal at the end of the day but, then again, we often have so many leftovers, they need to go in the bin. Especially the sides; sometimes we end up having whole trays of unsold food and you can't really tell what is going to sell well on a particular day.'

'The job pays well for the hours I have to work, but I need to have a second job in the evenings to make ends meet.'

'It's pretty boring in the morning and after 14.30.'

Questions

1. What do you think are the key issues faced by FreshLunch and what are the underlying reasons for these issues?
2. What advice would you give Sofia and how would you prioritise potential improvements?

Applying the principles

Some of these exercises can be answered by reading the chapter. Others will require some general knowledge of business activity and some might require an element of investigation. **Model answers for the first two questions are available on the student companion website**. *For model answers to the other questions in this section, please ask your tutor/lecturer.*

1. A pizza-making company has a demand forecast for the next 12 months that is shown in the table below. The current workforce of 100 staff can produce 1,500 cases of pizzas per month.

   a) Prepare a production plan that keeps the output level. How much warehouse space would the company need for this plan?

   b) Prepare a chase capacity plan. What implications would this have for staffing levels, assuming that the maximum amount of overtime would result in production levels of only 10 per cent greater than normal working hours?

Pizza demand forecast

Month	*Demand (cases per month)*
January	600
February	800
March	1,000
April	1,500
May	2,000
June	1,700
July	1,200
August	1,100
September	900
October	2,500
November	3,200
December	900

2. In a typical seven-day period, the planning department of the pizza company programs its ‘Pizzamatic’ machine for 148 hours. It knows that changeovers and set-ups take eight hours and breakdowns average four hours each week. Waiting for ingredients to be delivered usually accounts for six hours, during which time the machine cannot work. When the machine is running, it averages 87 per cent of its design speed. And inspection has revealed that 2 per cent of the pizzas processed by the machine are not up to the company’s quality standard. Calculate the OEE of the ‘Pizzamatic’ machine.

3. A German car manufacturer defines ‘utilisation’ as the ratio of actual output for a process to its design capacity, where design capacity is the capacity of a process as it is designed to operate. However, it knows that it is rarely possible to achieve this theoretical level of capacity, which is why the company uses a measure that it calls ‘effective capacity’. This is the actual capacity of a process, once maintenance, changeover, other stoppages and loading have been considered. The ratio of actual output for a process to its effective capacity is defined as its ‘efficiency’.

The company has a painting line with a design capacity of 100 square metres per minute and the line is operated 24 hours a day, 7 days a week (168 hours). Records for a week show the following lost time in production:

	Activity	*Number of hours lost*
1	Product changeovers (set-ups)	18 hrs
2	Regular maintenance	15 hrs
3	No work scheduled	6 hrs
4	Quality sampling checks	8 hrs
5	Shift change times	8 hrs
6	Maintenance breakdown	16 hrs
7	Quality failure investigation	12 hrs
8	Paint stock-outs	6 hrs
9	Labour shortages	6 hrs
10	Waiting for paint	5 hrs
	Total	**100 hrs**

During this week, production was only $100 \times 60 \times (168 - 100) = 408{,}000$ square metres per week. What is the painting line's 'utilisation' and 'efficiency' according to the company's definitions?

4. Seasonal demand is particularly important to the greetings card industry. Mother's Day, Father's Day, Halloween, Valentine's Day and other occasions have all been promoted as times to send (and buy) appropriately designed cards. Now, some card manufacturers have moved on to 'non-occasion' cards, which can be sent at any time. The cards include those intended to be sent from a parent to a child with messages such as 'Would a hug help?', 'Sorry I made you feel bad' or 'You're perfectly wonderful – it's your room that's a mess'. Other cards deal with more serious adult themes such as friendship ('You're more than a friend, you're just like family') or even alcoholism ('This is hard to say, but I think you're a much neater person when you're not drinking'). Some card companies have founded 'loyalty marketing groups' that '*help companies communicate with their customers at an emotional level*'. They promote the use of greetings cards for corporate use, to show that customers and employees are valued.

a) What seem to be the advantages and disadvantages of these strategies?

b) What else could card companies do to cope with demand fluctuations?

5. Panettone is *the* Italian Christmas cake, traditionally made in Milan. Now it has become popular outside its traditional Italian markets, and more than 40 million of them are consumed throughout Italy and all over the world during the holiday period. This boost in demand is good news for the big Italian manufacturers, but although volumes are higher, the product is still highly seasonal, which poses a problem for even the experienced Milanese confectioners. What could panettone makers do to cope with these demand fluctuations?

Notes on chapter

1 The information on which this example is based was taken from: Adams, S. (2020) 'Online learning platform Udemy said to seek funding at $3 billion valuation', *Forbes*, 15 July; *The New York Times* (2018) 'What will New York do about its Uber problem?', 7 May; Vora, S. (2018)' From new Airbnb Plus, properties with more amenities', *The New York Times*, 22 February; https://www.uber.com/gb/en/; https://www.udemy.com/; https://www.airbnb.co.uk/ [all accessed 7 October 2020].

2 The information on which this example is based was taken from: Heaven, W. D. (2020) 'Our weird behavior during the pandemic is messing with AI models', *MIT Technology Review*, 11 May; Kumar, N. and Haydon, D. (2020) 'Industries most and least impacted by COVID-19 from a probability of default perspective', S&P Global, 7 April; McKinsey & Company (2020) 'COVID-19: Global health and crisis response', 6 July.

3 The information on which this example is based was taken from: Gadher, D. (2019) 'Art-lovers see red at surge pricing', *The Sunday Times*, 18 August; Economist (2016) 'Jacking up prices may not be the only way to balance supply and demand for taxis', *Economist* print edition, 14 May; Dholakia, U.M. (2015) 'Everyone hates Uber's surge pricing – here's how to fix it', *Harvard Business Review*, 21 December.

4 The information on which this example is based was taken from: Das, A.K. (2019) 'Six bidders vie for Indian Railways ETCS Level 2 pilot project', *International Railway Journal*, 7 November; Rail Technology Magazine (2017) 'Network Rail awards landmark €150m ETCS signalling contract', 20 December; Railway Pro (2018) 'India to install ETCS Level 2 on its entire broad-gauge network', 8 March; Jha, S. (2018) 'Modi blocks Indian Railways ETCS plan', *International Railway Journal*, 11 April.

5 The information on which this example is based is taken from: Chong, A. (2019) 'What will it take for LTA's latest anti-congestion plan to work?', *Channel News Asia*, International Edition, 13 May; Economist (2015) 'What the London Underground reveals about work in the capital', *Economist* print edition, 23 May.

6 The information on which this example is based is taken from: Paton, G. (2017) 'Airline price war means you could be dragged off your next flight', *The Times*, 11 April; BBC (2017) 'United CEO says removed passenger was "disruptive and belligerent"', BBC News website, 11 April; Gunter, J. (2017) 'United Airlines incident: Why do airlines overbook?', BBC News, 10 April; Hill, A. (2017) 'United's reputational repair job mixes sense and nonsense', *Financial Times*, 27 April.

Taking it further

Gilliland, M., Tashman, L. and Sglavo, U. ***(eds) (2016) Business Forecasting: Practical problems and solutions, John Wiley & Sons.*** *A collection of papers focused on forecasting practitioners.*

Gunther, N.J. ***(2007) Guerrilla Capacity Planning, Springer.*** *This book provides a tactical approach for planning capacity in both product-based and service-based contexts. Particularly interesting for those new to the ideas of capacity planning, as it covers basic and more advanced demand forecasting techniques as well as 'classic' capacity responses.*

Kolassa, S. and Siemsen, E. ***(2016) Demand Forecasting for Managers, Business Expert Press.*** *Exactly what it says.*

Manas, J. ***(2014) The Resource Management and Capacity Planning Handbook: A guide to maximizing the value of your limited people resources, McGraw-Hill Education.*** *A practitioner's guide, particularly focused on managing human resource capacity to deliver better performance.*

Ord, K., Fildes, R. and Kourentzes, N. ***(2017) Principles of Business Forecasting, 2nd edition, Wessex, Inc.*** *A very detailed textbook covering demand in real depth.*

9 Inventory management

Introduction

Operations managers often have an ambivalent attitude towards inventories – of material, of people (queues), or of information. They can be costly, tying up working capital, they can annoy customers by making them wait, they are risky because items or information held in stock could deteriorate, become obsolete or just get lost, and they also can take up valuable space in the operation. On the other hand, inventories can provide some security in an uncertain environment. Knowing that you have the things, people or information 'in stock' is a comforting insurance against demand fluctuations. This is the dilemma of inventory management: in spite of the cost and the other disadvantages associated with holding stocks, they do facilitate the smoothing of supply and demand. In fact, they only exist because supply and demand are not exactly in harmony with each other. Figure 9.1 shows the position of the ideas described in this chapter within the general model of operations management.

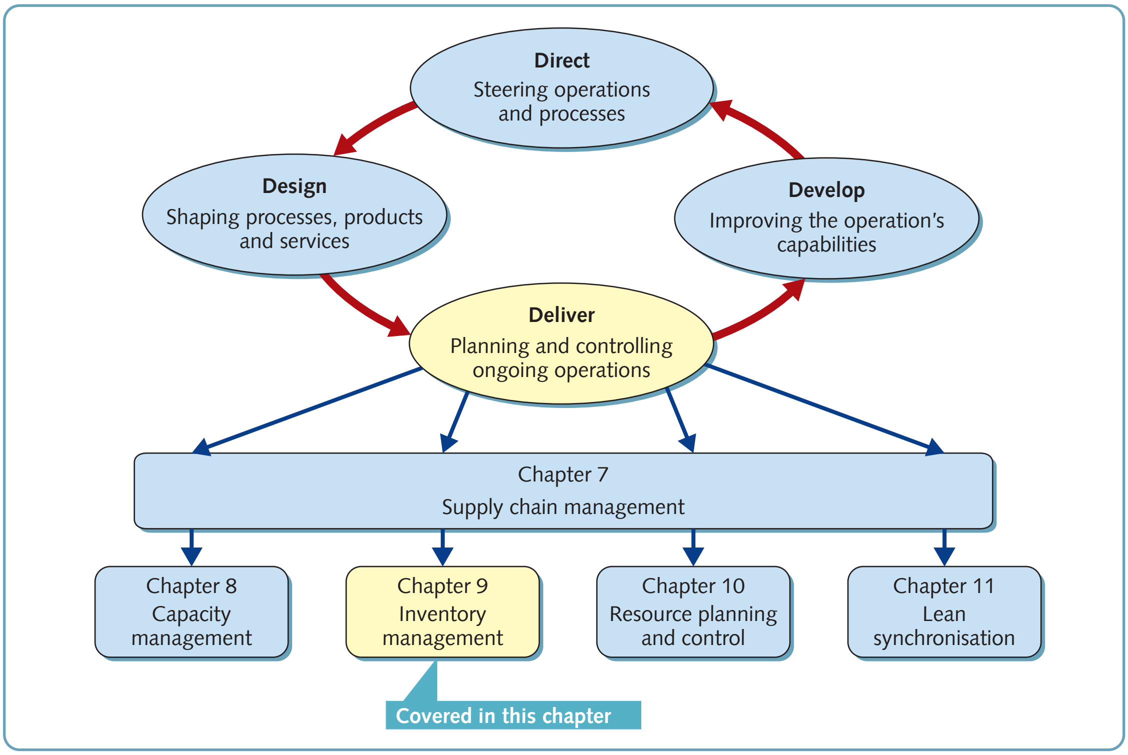


Figure 9.1 **Inventory management is the activity of planning and controlling accumulations of transformed resources as they move through supply networks, operations and processes**

EXECUTIVE SUMMARY

9.1 Is the role of inventory understood?

↓

9.2 Why should there be any inventory?

↓

9.3 Is the right quantity being ordered?

↓

9.4 Are inventory orders being placed at the right time?

↓

9.5 Is inventory being controlled effectively?

9.1 Is the role of inventory understood?

Inventory management is the activity of planning and controlling accumulations of transformed resources as they move through supply networks, operations and processes. The inventory can be accumulations of materials, customers or information. Accumulations of inventory occur because of local mismatches between supplier and demand. All operations have inventories of some kind and inventory management is particularly important where the inventories are central to the operation's objectives and/or of high value. How inventories are managed will determine the balance between customer service and cost objectives.

9.2 Why should there be any inventory?

Although there are cost, space, quality and operational/organisational disadvantages with inventory, there are also benefits. As far as physical inventory is concerned, it can act as an insurance against uncertainty, or a buffer against unexpected fluctuations in supply and demand, or to counteract a lack of flexibility, or to allow operations to take advantage of short-term opportunities, to anticipate future demands, to reduce overall costs, or even because it may increase in value. Inventories (queues) of customers can again help balance capacity and demand, or can enable prioritisation, or give customers time to choose, and enable efficient use of resources. Inventories of information (databases) can provide efficient multi-level access, allow single data capture and speed some processes. The underlying objective of inventory management is often to minimise inventory while maintaining acceptable customer service.

9.3 Is the right quantity being ordered?

A key inventory decision is the 'order quantity' decision. Various formulae exist that attempt to identify the order quantity that minimises total costs under different circumstances. One approach to this problem, the news vendor problem, includes the effects of probabilistic demand in determining order quantity.

9.4 Are inventory orders being placed at the right time?

Broadly, there are two approaches to this. The reorder point approach is to time reordering at the point in time where stock will fall to zero minus the order lead time. A variation of this is to reorder at the equivalent inventory level (the reorder level approach). Reordering at a fixed point or level are termed continuous review methods because they require continuous monitoring of stock levels. A different approach, called the periodic review approach, places orders at predetermined times, but varies the order depending on the level of inventory at that time. Both continuous and periodic review can be calculated on a probabilistic basis to include safety stocks.

9.5 Is inventory being controlled effectively?

The most common inventory control approach is based on the Pareto (at 80:20) curve. It classifies stocked items by the usage value (their usage rate multiplied by their value). High usage value items are deemed A class and controlled carefully, whereas low usage value items (B and C class) are controlled less intensely. However, this approach often has to be modified to take account of slow-moving items. Inventory information systems are generally used to keep track of inventory, forecast demand and place orders automatically.

9.1 Diagnostic question: *Is the role of inventory understood?*

Inventory is a term we use to describe the accumulations of materials, customers or information as they flow through processes or networks. Occasionally, the term is also used to describe transforming resources, such as rooms in hotels or automobiles in a vehicle hire firm, but here we use the term for the accumulation of resources that flow through processes, operations or supply networks. Physical inventory (sometimes called 'stock') is the accumulation of physical materials such as components, parts, finished goods or physical (paper) information records. Queues are accumulations of customers – physical as in a queuing line or people in an airport departure lounge, or waiting for service at the end of a phone line. Databases are stores for accumulations of digital information, such as medical records or insurance details. Managing these accumulations is what we call 'inventory management'. And it's important. Material inventories in a factory can represent a substantial proportion of cash tied up in working capital. Minimising them can release large quantities of cash. However, reducing them too far can lead to customers' orders not being fulfilled. Customers held up in queues for too long can get irritated, angry, and possibly leave, so reducing revenue. Databases are critical for storing digital information and while storage may be inexpensive, maintaining databases may not be.

> **OPERATIONS PRINCIPLE**
> ***All processes, operations and supply networks have inventories (accumulations) of materials, customers and information.***

Inventories are the result of uneven flow

Most things that flow do so in an uneven way. Rivers flow faster down steep sections or where they are squeezed into a ravine. Over relatively level ground they flow slowly, and form pools or even large lakes where there are natural or man-made barriers blocking their path. It's the same in operations. Passengers in an airport flow from public transport or their vehicles then have to queue at several points, including check-in, security screening and immigration. They then have to wait (in a queue even if they are sitting) again in the departure lounge as they are joined (batched) with other passengers to form a group of several hundred people who are ready to board the aircraft. They are then squeezed down the air bridge as they file in one at a time to board the plane. Likewise, in a tractor assembly plant, stocks of components such as gearboxes, wheels, lighting circuits, etc. are brought into the factory in 10s or 100s and are then stored next to the assembly line ready for use. Finished tractors will also be stored until the transporter comes to take them away in 1s or 10s to the dealers, or directly to the end customer. Similarly, a government tax department collects information about us and our finances from various sources, including our employers, our tax forms, information from banks or other investment companies, and stores this in databases until it is checked, sometimes by people, sometimes automatically, to create our tax codes and/or tax bills. In fact, because most operations involve flows of materials, customers and/or information, at some points they are likely to have material and information inventories and queues of customers waiting for goods or services (see Table 9.1).

Inventories are often the result of uneven flows. If there is a difference between the timing of supply and demand, or the rate of supply and demand, at any point in a process or network then accumulations will occur. A common analogy is the water tank shown in Figure 9.2. If, over time, the rate of supply of water to the tank differs from the rate at which it is demanded, a tank of water (inventory) will be needed to maintain supply. When the rate of supply exceeds the rate of demand, inventory increases; when the rate of demand exceeds the rate of supply, inventory decreases. So if an operation or process can match supply and demand rates, it will also succeed in reducing its inventory levels. But most organisations must cope with unequal

Table 9.1 Examples of inventory held in processes, operations or supply networks

Process, operation or supply network	*Inventories*		
	Physical inventories	*Queues of customers*	*Information in databases*
Hotel	Food items, drinks, toiletry items	Check-in and check-out	Customer details, loyalty card holders, catering suppliers
Hospital	Dressings, disposable instruments, blood	Patients on a waiting list, patients in bed waiting for surgery, patients in recovery wards	Patient medical records
Credit card application process	Blank cards, form letters	Customers waiting on the phone	Customers' credit and personal information
Computer manufacturer	Components for assembly, packaging materials, finished computers ready for sale	Customers waiting for delivery of their computer	Customers' details, supplier information

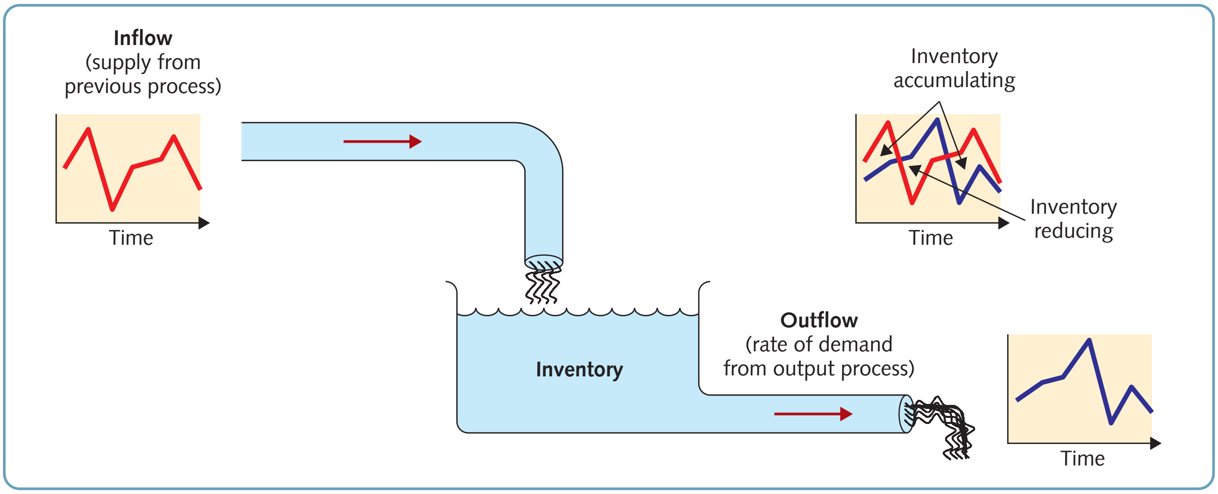


Figure 9.2 Inventory is created to compensate for the differences in timing between supply and demand

supply and demand, at least at some points in their supply chain. Organisations depend on the ability to manage this supply and demand inequality through their inventory management.

There is a complication when using the 'water flow' analogy to represent flows and accumulations (inventories) of information. Inventories of information can either be stored because of uneven flow, in the same way as materials and people, or stored because the operation needs to use the information to process something in the future. For example, an internet retail operation will process each order it receives, and inventories of information may accumulate because of uneven flows as we have described. But, in addition, during order processing customer details could be permanently stored in a database. This information will then be used, not only for future orders from the same customer, but also for other processes such as targeting promotional activities. In this case the inventory of information has turned from a transformed resource into a transforming resource, because it is being used to transform other information rather than being transformed itself. So, whereas managing physical material concerns ordering and holding the right amounts of goods or materials to deal with the variations in flow, and managing queues is about the level of resources to deal with demand, a database is the accumulation of information but it may not cause an interruption to the flow. Managing databases is about the organisation of the data, its storage, security and retrieval (access and search).

Cost and service – the inventory trade-off

Managing inventory can be complex and uncertain in practice, but at its heart it is a process of attempting to manage the trade-off between the costs of holding stock against the customer service that comes from having appropriate stock levels. Stock levels that are too high have a cost. This may largely be the working capital consequences in the case of, for example, road grit than can be stored outside, or it could also be the running costs of keeping the inventory in acceptable condition, such as in the case of frozen food, or it could also be the deterioration of the stocks, as in the case of fresh blood in a blood bank. Yet without an appropriate level of inventory, customers suffer poor service. This means potentially disrupting local traffic flow in the case of road grit, or failing to supply supermarkets with frozen food, with consequent lost profits. But a failure of supply from a blood bank may have even more drastic consequences. This type of trade-off is managed through a continuous stream of decisions at each point in the inventory system made every day by operations managers. Requests for grit, frozen food or blood are received from internal or external customers; they will be supplied, and demand will gradually deplete the inventory. Orders will need to be placed for replenishment of the stocks; deliveries will arrive and require storing. And at each stage of managing the inventory the competing demands of costs and service levels will need trading off. But to manage this trade-off we first need to understand the reasons for not having inventory, the reasons for having it, and then understand the tools available to help make these balancing decisions.

9.2 Diagnostic question: *Why should there be any inventory?*

There are plenty of reasons to avoid accumulating inventory where possible. Table 9.2 identifies some of these – particularly those concerned with cost, space, quality and operational/organisational issues.

So why have inventory?

On the face of it, it may seem sensible to have a smooth and even flow of materials, customers and information through operational processes and networks, and thus not have any accumulations. However, inventories provide many advantages for both operations and their customers. If a customer has to go to a competitor because a part is out of stock, or because they have had to wait too long, or because the company insists on collecting all their personal details each time they call, the value of inventories seems indisputable. The task of operations management is to allow inventory to accumulate only when its benefits outweigh its disadvantages. The following are some of the benefits of having an inventory.

> OPERATIONS PRINCIPLE
> ***Inventory should accumulate only when the advantages of having it outweigh its disadvantages.***

Physical inventory is an insurance against uncertainty

Inventory can act as a buffer against unexpected fluctuations in supply and demand. For example, a retail operation can never forecast demand perfectly over the lead time. It will order goods from its suppliers such that there is always a minimum level of inventory to cover against the possibility that demand will be greater than expected during the time taken to deliver the goods. This is buffer, or safety stock. It can also compensate for the uncertainties in

Table 9.2 Some reasons to avoid inventories

	Inventories		
	Physical inventories	***Queues of customers***	***Digital information in databases***
Cost	Ties up working capital and there could be high administrative and insurance costs	Primarily time-cost to the customer, i.e. wastes customers' time	Cost of set-up, access, updates and maintenance
Space	Requires storage space	Requires areas for waiting or phone lines for held calls	Requires memory capacity; may require secure and/or special environment
Quality	May deteriorate over time, become damaged or obsolete	May upset customers if they have to wait too long; may lose customers	Data may be corrupted or lost or become obsolete
Operational/ organisational	May hide problems (see lean synchronisation in Chapter 11)	May put undue pressure on staff and so quality is compromised for throughput	Databases need constant management, access control, updating and security

the process of the supply of goods into the store. The same applies with the output inventories, which is why hospitals always have a supply of blood, sutures and bandages for immediate response to accident and emergency patients. Similarly, auto-servicing services, factories and airlines may hold selected critical spare parts inventories so that maintenance staff can repair the most common faults without delay. Again, inventory is being used as an 'insurance' against unpredictable events.

Case example

Where was the safety stock when COVID-19 struck?[1]

The idea of safety stock is not new. Drastic and inherently unpredictable interruptions to supply have always been recognised by most operations managers – at least in theory. The problem is that what seems obvious when disaster strikes can seem like a waste of resources in normal times. Earthquakes in Japan or machinery breaking down in a Polish-based supplier's factory are the types of rare-but-inconvenient occurrences for which safety stocks are intended, and some operations will probably keep stocks to protect themselves. But what about global pandemics, with worldwide factory closures, one's own operations at least partially closed by staff 'lockdowns', delays or non-existent transportation links and, in some cases, heightened demand or even panic-buying? That is what healthcare operations faced at the start of the COVID-19 pandemic. With the global economy and international markets plunging as the coronavirus spread from China to other countries in Asia, Australia, Europe, the Americas and the Middle East, governments attempted to prevent the further spread of the virus by imposing travel restrictions and quarantines, further disrupting international supply chains. Countries varied in the amount and types of safety stocks they held at the start of the outbreak. Most health systems had recognised the possibility of a pandemic, even though it was seen as a 'black swan event' – an event that has a very high impact, but a very low probability of happening. And most kept some kind of safety stockpile. But often the safety stocks were designed to cover 'normal' fluctuations or spikes in demand, not the massive increase sparked by COVID-19. In the UK, for example, the stock of personal protective equipment (PPE) was a maximum of five or six weeks' worth, either in Britain or en route. Also, what constitutes the type of emergency that can release safety stock? In the UK, technically the stockpile could only be triggered when the World Health Organization declared a 'pandemic influenza', yet COVID-19 was not flu. The government body responsible had to intervene (relatively quickly) to order the release of the stockpile. Also, the items in the safety stockpile were designed for a flu pandemic, but COVID-19 was a different disease with a higher hospitalisation rate. So, it contained the surgical face masks, FFP3 respirators, gloves and aprons needed to tackle an influenza outbreak, but not sufficient fluid-repellent gowns and visors that were to prove critical for treating a novel virus like COVID-19, which could survive much longer outside the body. Nor can safety stock items usually be quickly replaced. That's why one keeps them as inventory. The panic to buy PPE at the beginning of the pandemic saw governments and private health organisations alike scrambling to replenish their stocks. States and companies were involved in a bidding war. Consignments of PPE were spirited away from airports before they could be sold to rivals. Some shipments were even diverted when the supplier received a higher bid.

American Photo Archive/Alamy Stock Photo

Physical inventory can counteract a lack of flexibility

Where a wide range of customer options is offered, unless the operation is perfectly flexible, stock will be needed to ensure supply when it is engaged on other activities. This is sometimes called cycle inventory. For example, suppose a baker makes three types of bread. Because of the nature of the mixing and baking process, only one kind of bread can be produced at any time. The baker will have to produce each type of bread in batches large enough to satisfy the demand for each kind of bread between the times when each batch is ready for sale. So, even when demand is steady and predictable, there will always be some inventory to compensate for the intermittent supply of each type of bread.

Physical inventory allows operations to take advantage of short-term opportunities

Sometimes opportunities arise that necessitate accumulating inventory, even when there is no immediate demand for it. For example, a supplier may be offering a particularly good deal on selected items for a limited time period, perhaps because they want to reduce their own finished goods inventories. Under these circumstances, a purchasing department may opportunistically take advantage of the short-term price advantage.

Physical inventory can be used to anticipate future demands

Medium-term capacity management (covered in Chapter 8) may use inventory to cope with demand capacity. Rather than trying to make a product (such as chocolate) only when it is needed, it is produced throughout the year ahead of demand and put into inventory until it is needed. This type of inventory is called anticipation inventory and is most commonly used when demand fluctuations are large but relatively predictable.

Physical inventory can reduce overall costs

Holding relatively large inventories may bring savings that are greater than the cost of holding the inventory. This may be when bulk-buying gets the lowest possible cost of inputs, or when large order quantities reduce both the number of orders placed and the associated costs of administration and material handling. This is the basis of the 'economic order quantity' (EOQ) approach that will be treated later in this chapter.

Physical inventory can increase in value

Sometimes the items held as inventory can increase in value and so become an investment. For example, dealers in fine wines are less reluctant to hold inventory than dealers in wine that does not get better with age. (However, it can be argued that keeping fine wines until they are at their peak is really part of the overall process rather than inventory as such.) A more obvious example is inventories of money. The many financial processes within most organisations will try to maximise the inventory of cash they hold because it is earning them interest.

Physical inventory fills the processing 'pipeline'

'Pipeline' inventory exists because transformed resources cannot be moved instantaneously between the point of supply and the point of demand. When a retail store places an order, its supplier will 'allocate' the stock to the retail store in its own warehouse, pack it, load it onto its truck, transport it to its destination and unload it into the retailer's inventory. From the time that stock is allocated (and therefore it is unavailable to any other customer) to the time it becomes available at the retail store, it is pipeline inventory. Especially in geographically dispersed supply networks, pipeline inventory can be substantial.

Queues of customers help balance capacity and demand

This is especially useful if the main service resource is expensive – for example, doctors, consultants, lawyers or expensive equipment such as CAT scans. By waiting a short time after their arrival, and creating a queue of customers, the service always has customers to process. This is also helpful where arrival times are less predictable, such as where an appointment system is not used or is not possible.

Queues of customers enable prioritisation

In cases where resources are fixed and customers are entering the system with different levels of priority, the formation of a queue allows the organisation to serve urgent customers while keeping other less-urgent ones waiting. In the UK it is not unusual to have to wait 3–4 hours for treatment in an accident and emergency ward, with more urgent cases 'jumping the queue' for treatment.

Queuing gives customers time to choose

Time spent in a queue gives customers time to decide what products/services they require. For example, customers waiting in a fast-food restaurant have time to look at the menu so that when they get to the counter they are ready to make their order without holding up the server.

Queues enable efficient use of resources

By allowing queues to form, customers can be batched together to make efficient use of operational resources. For example, a queue for an elevator makes better use of its capacity; in an airport, by calling certain customers to the gate, staff can load the aircraft more efficiently and quickly.

Databases provide efficient multi-level access

Databases are relatively cheap ways of storing information and providing many people with access, although there may be restrictions or different levels of access. The doctor's receptionist will be able to call up a patient's records to check their name and address and make an appointment. The doctor will then be able to call up the appointment and the patient's records. The pharmacist will be able to call up the patient's name and prescriptions and cross-check for other prescriptions and known allergies, and so on.

Databases of information allow single data capture

There is no need to capture data at every transaction with a customer or supplier, though checks may be required.

Databases of information speed the process

For example, Amazon will (if you agree to it) store your delivery address and credit card information so that future purchases can be made with a single click, making it fast and easy for the customer.

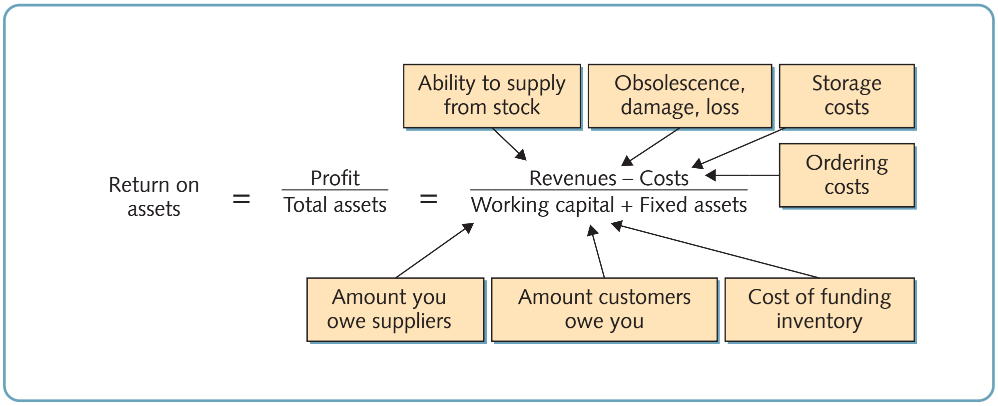


Figure 9.3 **Inventory management has a significant effect on return on assets**

The effect of inventory on return on assets

One can summarise the effects on the financial performance of an operation by looking at how some of the factors of inventory management impact on 'return on assets', a key financial performance measure. Figure 9.3 shows some of these factors:

- *Inventory governs the operation's ability to supply its customers*. An absence of inventory means that customers are not satisfied, with the possibility of reduced revenue.
- *Inventory may become obsolete* as alternatives become available, or could *be damaged, deteriorate or simply get lost*. This increases costs (because resources have been wasted) and reduces revenue (because the obsolete, damaged or lost items cannot be sold).
- *Inventory incurs storage costs* (leasing space, maintaining appropriate conditions, etc.). This could be high if items are hazardous to store (for example, flammable solvents, explosives, chemicals) or difficult to store, requiring special facilities (for example, frozen food).
- *Inventory involves administrative and insurance costs*. Every time a delivery is ordered, time and costs are incurred.
- *Inventory ties up money* in the form of working capital, which is therefore unavailable for other uses, such as reducing borrowings or making investment in productive fixed assets (we shall expand on the idea of working capital later).
- *Inventory contracts with suppliers can dictate the timing of when suppliers need to be paid*. If they require paying before the operation receives payment from its customers (as is normal), the difference between the amount the operation owes suppliers and the amount suppliers owe the operation adds to working capital requirements.

OPERATIONS PRINCIPLE
Inventory management can have a significant effect on return on assets.

Reducing physical inventory

The objective of most operations managers who manage physical inventories is to reduce the overall level (and/or cost) of inventory while maintaining an acceptable level of customer service. Table 9.3 identifies some of the ways in which inventory may be reduced.

Table 9.3 Some ways in which physical inventory may be reduced

Reason for holding inventory	*Example*	*How inventory could be reduced*
As an insurance against uncertainty	Safety stocks for when demand or supply is not perfectly predictable	• Improve demand forecasting • Tighten supply, e.g. through service-level penalties
To counteract a lack of flexibility	Cycle stock to maintain supply when other products are being made	• Increase flexibility of processes, e.g. by reducing change-over times (see Chapter 11) • Using parallel processes producing output simultaneously (see Chapter 5)
To take advantage of relatively short-term opportunities	Suppliers offer 'time-limited' special low-cost offers	• Persuade suppliers to adopt 'everyday low prices' (see Chapter 11)
To anticipate future demands	Build up stocks in low-demand periods for use in high-demand periods	• Increase volume flexibility by moving towards a 'chase demand' plan (see Chapter 8)
To reduce overall costs	Purchasing a batch of products in order to save delivery and administration costs	• Reduce administration costs through purchasing process efficiency gains • Investigate alternative delivery channels that reduce transport costs
To fill the processing 'pipeline'	Items being delivered to customer	• Reduce process time between customer request and dispatch of items • Reduce throughput time in the downstream supply chain (see Chapter 7)

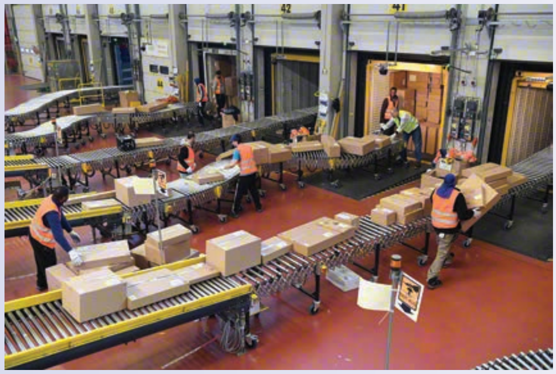
Case example

Amazon's 'anticipatory inventory'[2]

Could an item be on its way from a stocking point to you, even before you think of ordering it? Is it possible for a company to accurately forecast your order and ship it to you before you place it? Of course, forecast accuracy and time to deliver are related. Poor forecasts mean that the wrong items will be stored, which in turn means that delivery will be delayed until the right items are received. But what if a supplier could know what its customers were going to order, even before they do? That is the ambition of Amazon's online retail operation. It filed a patent to protect its system for the technology that hopes to predict what its customers will buy, even before they have clicked the 'order' button. The company, which is the world's largest online retailer, calls its new system 'anticipatory shipping' and perceives it as a way to speed up its delivery times. Amazon's patent application reveals the thinking behind the system. Its application says that one substantial disadvantage to the virtual storefront model is that in many instances, customers cannot receive their merchandise immediately upon purchase, but must instead wait for the product to be shipped to them. The availability of expedited shipping methods from various

▶

common carriers may to some extent compensate for the delay in shipment, but often at substantial additional cost that may rival the price paid for the merchandise. Such delays may dissuade customers from buying items from online merchants, particularly if those items are more readily available locally. The approach is reported as using several elements to predict what purchases a person may make. Factors to be taken into account could include age, income, previously purchased items, searched-for items, 'wish lists' and maybe even the time a user's cursor lingers over a product. Armed with this information, Amazon could ship items that are likely to be ordered to the inventory 'hub' nearest to the customer. So, when a customer really does order the item, it can be delivered far faster.



Uli Deck/dpa picture alliance/Alamy Stock Photo

9.3 Diagnostic question: *Is the right quantity being ordered?*

To illustrate this decision, consider how we manage our domestic inventory. We implicitly make decisions on order quantity – that is, how much to purchase at one time – by balancing two sets of costs: the costs associated with going out to purchase the food items and the costs associated with holding the stocks. The option of holding very little or no inventory of food and purchasing each item only when it is needed requires small amounts of money each time, but involves buying several times a day, which is inconvenient. Conversely, making one journey to the local superstore every few months and purchasing all the provisions we would need until our next visit reduces purchasing time and costs but requires a very large amount of money each time the trip is made – money that could otherwise be in the bank and earning interest. We might also have to invest in extra cupboard units and a very large freezer. Somewhere between these extremes lies an ordering strategy that will minimise the total costs and effort involved in purchasing food.

Inventory costs

A similar range of costs apply in commercial order-quantity decisions as in the domestic situation. These are the costs of placing an order, including preparing the documentation, arranging for the delivery to be made, arranging to pay the supplier for the delivery and the general costs of keeping all the information that allows us to do this. An 'internal order' on processes within an operation has equivalent costs. Price-discount costs for large orders or extra costs for small orders may also influence how much to purchase. If inventory cannot supply demand, there will be costs incurred by failing to supply customers. External customers may even take their business elsewhere. Internal stock-outs could lead to idle time at the next process, inefficiencies and eventually, again, dissatisfied external customers. There are the working capital costs of funding the lag between paying suppliers and receiving payment from customers. Storage costs are the costs associated with physically storing goods, such as renting, heating and lighting a warehouse, as well as insuring the inventory. While stored as inventory, there is a risk of obsolescence costs if the inventory is superseded (in the case of a change in fashion) or deteriorates with age (in the case of most foodstuffs).

Some of these costs will decrease as order size is increased; the first three costs (cost of placing an order, price-discount costs and stock-out costs) are like this. The other costs (working capital, storage and obsolescence costs) generally increase as order size is increased. But it may not be the same organisation that incurs each cost. For example, sometimes suppliers agree to hold consignment stock. This means that they deliver large quantities of inventory to their customers for them to store but will only charge for the goods as and when they are used. In the meantime, they remain the supplier's property so do not have to be financed by the customer, who does however provide storage facilities.

Inventory profiles

An inventory profile is a visual representation of the inventory level over time. Figure 9.4 shows a simplified inventory profile for one particular stock item in a retail operation. Every time an order is placed, Q items are ordered. The replenishment order arrives in one batch instantaneously. Demand for the item is then steady and perfectly predictable at a rate of D units per month. When demand has depleted the stock of the items entirely, another order of Q items arrives instantaneously, and so on. Under these circumstances:

The average inventory $= \frac{Q}{2}$ (because the two shaded areas in Figure 9.4 are equal)

The time interval between deliveries $= \frac{Q}{D}$

The frequency of deliveries = the reciprocal of the time interval $= \frac{D}{Q}$

The economic order quantity (EOQ) formula

The economic order quantity (EOQ) approach attempts to find the best balance between the advantages and disadvantages of holding stock. For example, Figure 9.5 shows two alternative order-quantity policies for an item. Plan A, represented by the blue line, involves ordering in quantities of 400 at a time. Demand in this case is running at 1,000 units per year. Plan B, represented by the red line, uses smaller but more frequent replenishment orders. This time only 100 are ordered at a time, with orders being placed four times as often. However, the average inventory for plan B is one-quarter of that for plan A.

To find out whether either of these plans, or some other plan, minimises the total cost of stocking the item, we need some further information, namely the total cost of holding one unit in stock for a period of time (C_h) and the total costs of placing an order (C_o).

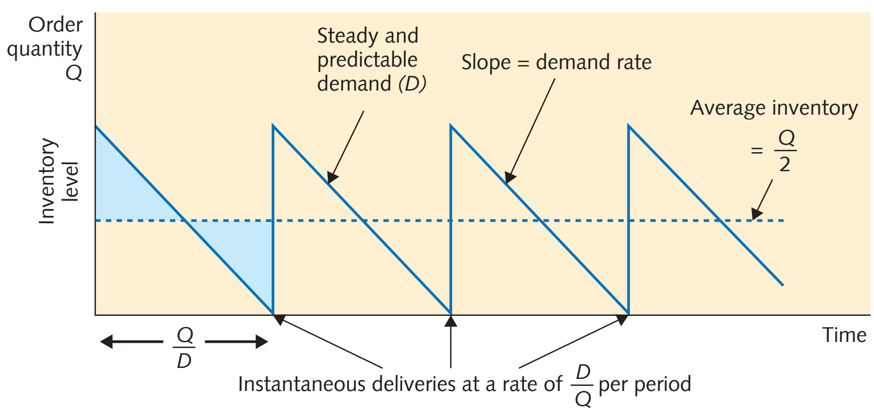


Figure 9.4 Inventory profiles chart the variation in inventory level

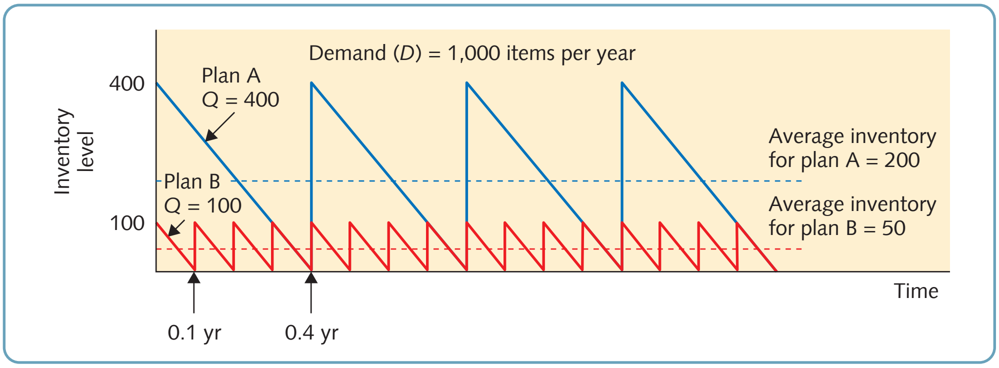


Figure 9.5 Two alternative inventory plans with different order quantities (*Q*)

In this case, the cost of holding stocks is calculated at £1 per item per year and the cost of placing an order is calculated at £20 per order.

We can now calculate total holding costs and ordering costs for any particular ordering plan as follows:

$$\text{Holding costs} = \text{holding cost/unit} \times \text{average inventory}$$

$$= C_h \times \frac{Q}{2}$$

$$\text{Ordering costs} = \text{ordering cost} \times \text{number of orders per period}$$

$$= C_0 \times \frac{D}{Q}$$

$$\text{Total cost, } C_t = \frac{C_h Q}{2} + \frac{C_0 D}{Q}$$

We can now calculate the costs of adopting plans with different order quantities. These are illustrated in Table 9.4. As we would expect, with low values of Q holding costs are low but ordering costs are high, because orders have to be placed very frequently. As Q increases, the holding costs increase but the costs of placing orders decrease. In this case, the order quantity, Q, which minimises the sum of holding and order costs, is 200. This 'optimum' order quantity is called the economic order quantity (EOQ). This is illustrated graphically in Figure 9.6.

Table 9.4 Costs of adoption of plans with different order quantities

Demand (*D*) = 1,000 units per year — Holding costs (C_h) = £1 per item per year
Order costs (C_o) = £20 per order

Order quantity (Q)	*Holding costs* ($0.5Q \times C_h$)	+	*Order costs* ($(D/Q) \times C_o$)	=	*Total costs*
50	25		20 × 20 = 400		425
100	50		10 × 20 = 200		250
150	75		6.7 × 20 = 134		209
200	100		5 × 20 = 100		200*
250	125		4 × 20 = 80		205
300	150		3.3 × 20 = 66		216
350	175		2.9 × 20 = 58		233
400	200		2.5 × 20 = 50		250

* Minimum total cost

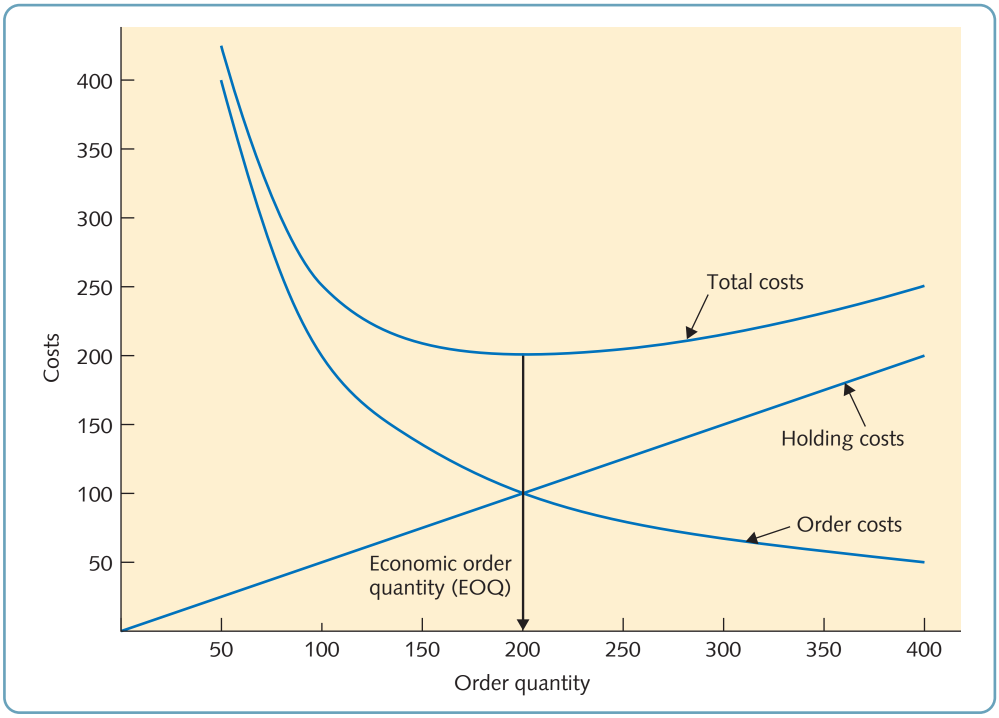


Figure 9.6 Inventory-related costs minimise at the 'economic order quantity' (EOQ)

A more elegant method of finding the EOQ is to derive its general expression. This can be done using simple differential calculus as follows. From before:

$$\text{Total cost} = \text{holding cost} + \text{order cost}$$

$$C_t = \frac{C_h Q}{2} + \frac{C_o D}{Q}$$

The rate of change of total cost is given by the first differential of C_t with respect to Q:

$$\frac{dC_t}{dQ} = \frac{C_h}{2} - \frac{C_o D}{Q^2}$$

The lowest cost will occur when $\frac{dC_t}{dQ} = 0$ that is:

$$0 = \frac{C_h}{2} - \frac{C_o D}{Q_o^2}$$

where $Q_o =$ the EOQ. Rearranging this expression gives:

$$Q_o = \text{EOQ} = \sqrt{\frac{2C_o D}{C_h}}$$

When using the EOQ:

$$\text{Time between orders} = \frac{\text{EOQ}}{D}$$

$$\text{Order frequency} = \frac{D}{\text{EOQ}} \text{ per period}$$

Sensitivity of the EOQ

The graphical representation of the total cost curve in Figure 9.6 shows that, although there is a single value of Q that minimises total costs, any relatively small deviation from the EOQ will

OPERATIONS PRINCIPLE
For any stock replenishment activity there is a theoretical 'optimum' order quantity (EOQ) that minimises total inventory-related costs.

not increase total costs significantly. In other words, costs will be near-optimum, provided a value of Q that is reasonably close to the EOQ is chosen. Put another way, small errors in estimating either holding costs or order costs will not result in a significant change in the EOQ. This is a particularly convenient phenomenon because, in practice, both holding and order costs are not easy to estimate accurately. The other implication is that, because the total cost curve is not symmetrical, it is usually better to have slightly more than slightly less inventory.

Gradual replacement – the economic batch quantity (EBQ) model

The simple inventory profile shown in Figure 9.4 assumes that each complete replacement order arrives at one point in time. However, replenishment may occur over a time period rather than in one lot – for example where an internal order is placed for a batch of parts to be produced on a machine. The machine will start to produce items and ship them in a more or less continuous stream into inventory, but at the same time demand is removing items from the inventory. Provided the rate at which items are being supplied to the inventory (P) is higher than the demand rate (D), then the inventory will increase. After the batch has been completed the machine will be reset (to produce some other part), and demand will continue to deplete the inventory level until production of the next batch begins. The resulting profile is shown in Figure 9.7. This is typical for inventories supplied by batch processes, and the minimum-cost batch quantity for this profile is called the economic batch quantity (EBQ). It is derived as follows:

$$\text{Maximum stock level} = M$$
$$\text{Slope of inventory build-up} = P - D$$

Worked example

CVM building materials

CVM building materials obtains its bagged cement from a single supplier. Demand is reasonably constant throughout the year, and last year the company sold 2,000 tonnes of this product. It estimates the costs of placing an order at around £25 each time an order is placed, and calculates that the annual cost of holding inventory is 20 per cent of purchase cost. The company purchases the cement at £60 per tonne. How much should the company order at a time?

$$\text{EOQ for cement} = \sqrt{\frac{2C_oD}{C_h}}$$
$$= \sqrt{\frac{2 \times 25 \times 2000}{0.2 \times 60}}$$
$$= \sqrt{\frac{100{,}000}{12}}$$
$$= 91.287 \text{ tonnes}$$

After calculating the EOQ, the operations manager feels that placing an order for 91.287 tonnes exactly seems somewhat over-precise. Why not order a convenient 100 tonnes?

Total cost of ordering plan for $Q = 91.287$:

$$= \frac{C_hQ}{2} + \frac{C_oD}{Q}$$
$$= \frac{(0.2 \times 60) \times 91.287}{2} + \frac{25 \times 2000}{91.287}$$
$$= £1095.454$$

Total cost of ordering plan for $Q = 100$:

$$= \frac{(0.2 \times 60) \times 100}{2} + \frac{25 \times 2000}{100}$$
$$= £1100$$

The extra cost of ordering 100 tonnes at a time is £1,100 − £1,095.45 = £4.55. The operations manager therefore should feel confident in using the more convenient order quantity.

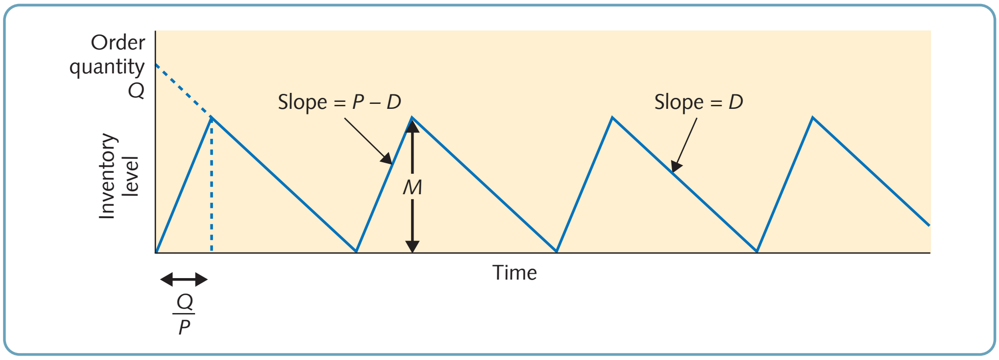


Figure 9.7 Inventory profile for gradual replacement of inventory

Also, as is clear from Figure 9.7:

$$\text{Slope of inventory build-up} = M \div \frac{Q}{P} = \frac{MP}{Q}$$

So,

$$\frac{MP}{Q} = P - D$$

$$M = \frac{Q(P - D)}{P}$$

$$\text{Average inventory level} = \frac{M}{2} = \frac{Q(P - D)}{2P}$$

As before:

$$\text{Total cost} = \text{holding cost} + \text{order cost}$$

$$C_t = \frac{C_h Q(P - D)}{2P} + \frac{C_o D}{Q}$$

$$\frac{dC_t}{dQ} = \frac{C_h(P - D)}{2P} - \frac{C_o D}{Q^2}$$

Again, equating to zero and solving Q gives the minimum-cost order quantity EBQ:

$$\text{EBQ} = \sqrt{\frac{2C_o D}{C_h(1 - (D/P))}}$$

If customers won't wait – the 'newsvendor problem'

A special case of the inventory order-quantity decision is when an order quantity is purchased for a specific event or time period after which the items are unlikely to be sold. A simple example of this is the decision taken by a newspaper vendor of how many newspapers to stock for the day. If the newsvendor should run out of papers, customers will either go elsewhere or decide not to buy a paper that day. Newspapers left over at the end of the day are worthless, and demand for the newspapers varies day by day. In deciding how many newspapers to carry, the newsvendor is in effect balancing the risk and consequence of running out of newspapers

Worked example

Clonacola

The manager of a Clonacola bottle-filling plant, which bottles soft drinks, needs to decide how long a 'run' of each type of drink to process. Demand for each type of drink is reasonably constant at 80,000 per month (a month has 160 production hours). The bottling lines fill at a rate of 3,000 bottles per hour, but take an hour to clean and reset between different drinks. The cost (of labour and lost production capacity) of each of these changeovers has been calculated at £100 per hour. Stock-holding costs are counted at £0.1 per bottle per month.

$$D = 80{,}000 \text{ per month}$$

$$= 500 \text{ per hour}$$

$$\text{EBQ} = \sqrt{\frac{2C_oD}{C_h(1-(D/P))}}$$

$$= \sqrt{\frac{2 \times 100 \times 80000}{0.1(1-(500/3000))}}$$

$$\text{EBQ} = 13{,}856$$

The staff who operate the lines have devised a method of reducing the changeover time from 1 hour to 30 minutes. How would that change the EBQ?

$$\text{New } C_o = £50$$

$$\text{New EBQ} = \sqrt{\frac{2 \times 50 \times 80000}{0.1(1-(500/3000))}}$$

$$= 9{,}798$$

against that of having newspapers left over at the end of the day. Retailers and manufacturers of high-class leisure products, such as some books and popular music CDs, face the same problem. For example, a concert promoter needs to decide how many concert T-shirts emblazoned with the logo of the main act to order. The profit on each T-shirt sold at the concert is £5 and any unsold T-shirts are returned to the company that supplies them, but at a loss to the promoter of £3 per T-shirt. Demand is uncertain but is estimated to be between 200 and 1,000. The probabilities of different demand are as follows:

Demand level	200	400	600	800
Probability	0.2	0.3	0.4	0.1

How many T-shirts should the promoter order? Table 9.5 shows the profit that the promoter would make for different order quantities and different levels of demand.

We can now calculate the expected profit that the promoter will make for each order quantity by weighting the outcomes by their probability of occurring.

If the promoter orders 200 T-shirts:

$$\text{Expected profit} = 1000 \times 0.2 + 1000 \times 0.3 + 1000 \times 0.4 + 1000 \times 0.1$$

$$= £1{,}000$$

If the promoter orders 400 T-shirts:

$$\text{Expected profit} = 400 \times 0.2 + 2000 \times 0.3 + 2000 \times 0.4 + 2000 \times 0.1$$

$$= £1{,}680$$

If the promoter orders 600 T-shirts:

$$\text{Expected profit} = -200 \times 0.2 + 1400 \times 0.3 + 3000 \times 0.4 + 3000 \times 0.1$$

$$= £1{,}880$$

Table 9.5 Pay-off matrix for T-shirt order quantity (profit or loss in £s)

Demand level	*200*	*400*	*600*	*800*
Probability	0.2	0.3	0.4	0.1
Promoter orders 200	1,000	1,000	1,000	1,000
Promoter orders 400	400	2,000	2,000	2,000
Promoter orders 600	−200	1,400	3,000	3,000
Promoter orders 800	−800	800	2,400	4,000

If the promoter orders 800 T-shirts:

$$\text{Expected profit} = -800 \times 0.2 + 800 \times 0.3 + 2400 \times 0.4 + 4000 \times 0.1$$
$$= £1{,}440$$

The order quantity that gives the maximum profit is 600 T-shirts, which results in a profit of £1,880.

The importance of this approach lies in the way it takes a probabilistic view of part of the inventory calculation demand – something we shall use again in this chapter.

Case example

Mr Ruben's bakery[3]

Be careful about treating the newsvendor problem on a product-by-product basis. It is a powerful idea, but needs to be seen in context. Take the famous City Bakery, in Manhattan, New York. It is run by Maury Rubin, a master baker, who knows the economics of baking fresh products. Ingredients and rent are expensive. It costs Mr Rubin $2.60 to make a $3.50 croissant. If he makes 100 and sells 70, he earns $245 but his costs are $260, and because all goods are sold within a day (his quality standards mean that he won't sell leftovers), he loses money. Nor can he raise his prices. In his competitive market, he says, shoppers bristle when the cost of baked goods passes a certain threshold. However, Mr Ruben has two 'solutions'. First, he can subsidise his croissants by selling higher-margin items such as fancy salads and sandwiches. Second, he uses data to cut waste, by studying sales so that he can detect demand trends in order to fine-tune supply. He monitors the weather carefully (demand drops away when it rains) and carefully inspects school calendars so he can reduce the quantities he bakes during school holidays. Each morning, he makes sure that pastries are prepared, but then he checks sales every 60–90 minutes before making the decision whether to adjust supply or not. Only when the numbers are in do the pastries go into the oven. Having no croissants left by the end of the day is a sign of success.

9.4 Diagnostic question: *Are inventory orders being placed at the right time?*

When we assumed that orders arrived instantaneously, and demand was steady and predictable, the decision on when to place a replenishment order was self-evident. An order would be placed as soon as the stock level reached zero, it would arrive instantaneously and prevent any stock-out occurring. When there is a lag between the order being placed and it arriving in the inventory, we can still calculate the timing of a replacement order simply, as shown in Figure 9.8. The lead time for an order to arrive is in this case two weeks, so the reorder point (ROP) is

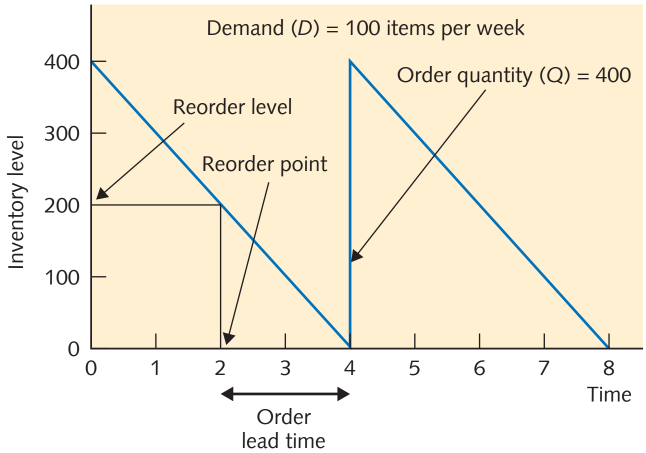


Figure 9.8 Reorder level (ROL) and Reorder point (ROP) are derived from the order lead time and demand rate

the point at which stock will fall to zero minus the order lead time. Alternatively, we can define the point in terms of the level that the inventory will have reached when a replenishment order needs to be placed. Here, this occurs at a reorder level (ROL) of 200 items.

However, this assumes that both the demand and the order lead time are perfectly predictable. In most cases this is not so. Both demand and order lead time are likely to vary to produce a profile that looks something like that in Figure 9.9. In these circumstances it is necessary to make the replenishment order somewhat earlier than would be the case in a purely deterministic situation. This will result in, on average, some 'safety' stock still being in the inventory when the replenishment order arrives. The earlier the replenishment order is placed, the higher will be the expected level of safety stock when the replenishment order arrives. But because of the variability of both lead time (t) and demand rate (d), the safety stock at the time of replenishment will vary. The main consideration in setting safety stock is the probability that the stock will not have run out before the replenishment order arrives. This depends on the lead-time usage distribution. This is a combination of the distributions that describe lead-time variation and the demand rate during the lead time. If safety stock is set below the lower limit of this distribution then there will be shortages every single replenishment cycle. If safety stock is set

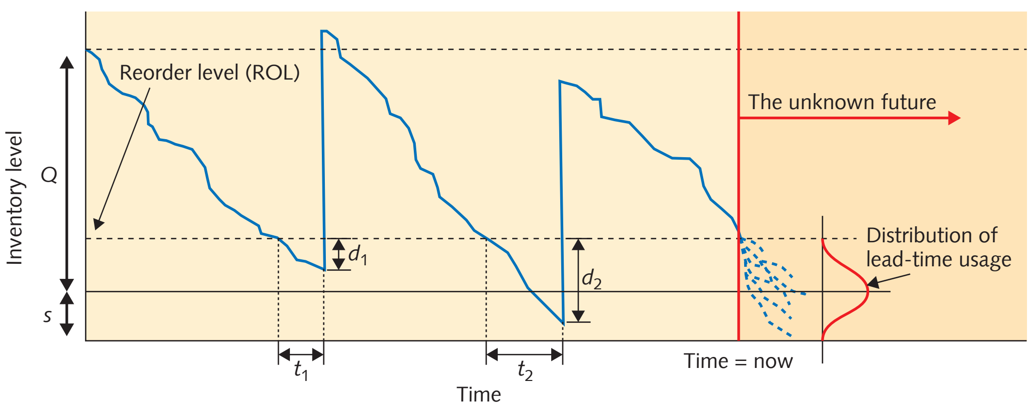


Figure 9.9 Safety stock (*s*) helps to avoid stock-outs when demand and/or order lead time are uncertain

OPERATIONS PRINCIPLE

For any stock replenishment activity, the timing of replenishment should reflect the effects of uncertain lead time and uncertain demand during that lead time.

above the upper limit of the distribution, there is no chance of stock-outs occurring. Usually, safety stock is set to give a predetermined likelihood that stock-outs will not occur. Figure 9.9 shows that, in this case, the first replenishment order arrived after t_1, resulting in a lead-time usage of $d1$. The second replenishment order took longer, t_2, and demand rate was also higher, resulting in a lead-time usage of d_2. The third order cycle shows several possible inventory profiles for different conditions of lead-time usage and demand rate.

Worked example

Knacko running shoes

Knacko, an online retailer of running shoes, can never be certain of how long, after placing an order, the delivery will take. Examination of previous orders reveals that out of 10 orders: one took one week, two took two weeks, four took three weeks, two took four weeks and one took five weeks. The rate of demand for the shoes also varies between 110 pairs per week and 140 pairs per week. There is a 0.2 probability of the demand rate being either 110 or 140 pairs per week, and a 0.3 chance of demand being either 120 or 130 pairs per week. The company needs to decide when it should place replenishment orders if the probability of a stock-out is to be less than 10 per cent.

Both lead time and the demand rate during the lead time will contribute to the lead-time usage. So the distributions that describe each will need to be combined. Figure 9.10 and Table 9.6 show how this can be done. Taking lead time to be either one, two, three, four or five weeks, and demand rate to be either 110, 120, 130 or 140 pairs per week, and also assuming the two variables to be independent, the distributions can be combined as shown in

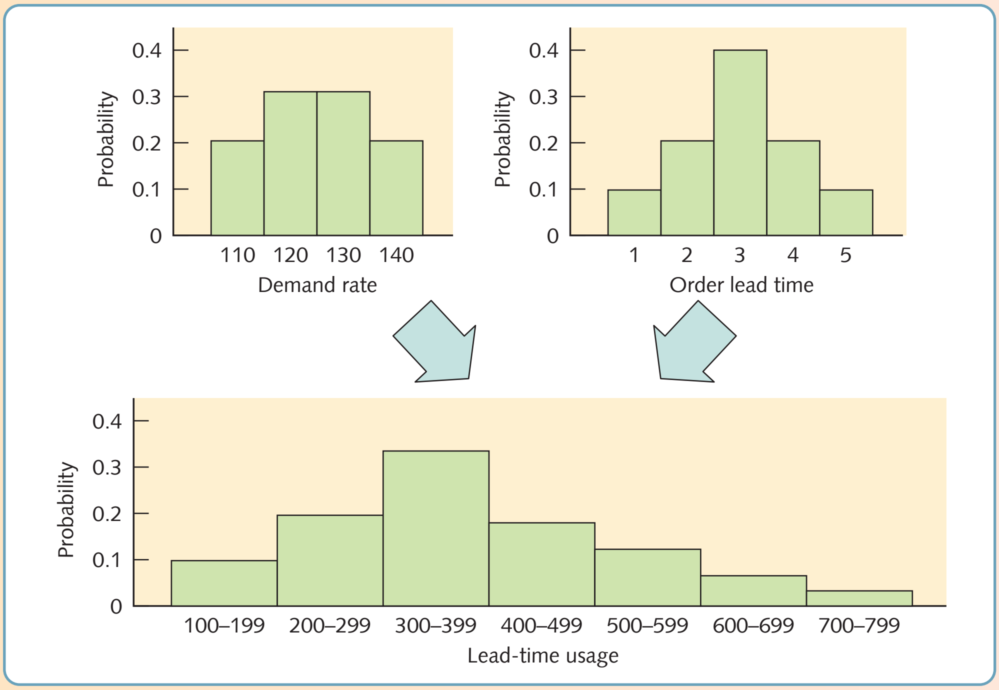


Figure 9.10 The probability distributions for order lead time and demand rate combine to give the lead-time usage distribution

Table 9.6. Each element in the matrix shows a possible lead-time usage with the probability of its occurrence. So, if the lead time is one week and the demand rate is 110 pairs per week, the actual lead-time usage will be $1 \times 110 = 110$ pairs. Since there is a 0.1 chance of the lead time being one week, and a 0.2 chance of demand rate being 110 pairs per week, the probability of both these events occurring is $0.1 \times 0.2 = 0.02$.

We can now classify the possible lead-time usages into histogram form. For example, summing the probabilities of all the lead-time usages that fall within the range 100–199 (all the first column) gives a combined probability of 0.1. Repeating this for subsequent intervals results in Table 9.7.

This shows the probability of each possible range of lead-time usage occurring, but it is the cumulative probabilities that are needed to predict the likelihood of stock-out (see Table 9.8).

Setting the reorder level at 600 would mean that there is only a 0.08 chance of usage being greater than available inventory during the lead time, i.e. there is a less-than-10 per cent chance of a stock-out occurring.

Table 9.6 Matrix of lead-time and demand-rate probabilities

			Lead-time probabilities				
			1	***2***	***3***	***4***	***5***
			0.1	***0.2***	***0.4***	***0.2***	***0.1***
Demand-rate probabilities	110	0.2	110 (0.02)	220 (0.04)	330 (0.08)	440 (0.04)	550 (0.02)
	120	0.3	120 (0.03)	240 (0.06)	360 (0.12)	480 (0.06)	600 (0.03)
	130	0.3	130 (0.03)	260 (0.06)	390 (0.12)	520 (0.06)	650 (0.03)
	140	0.2	140 (0.02)	280 (0.04)	420 (0.08)	560 (0.04)	700 (0.02)

Table 9.7 Combined probabilities

Lead-time usage	*100–199*	*200–299*	*300–399*	*400–499*	*500–599*	*600–699*	*700–799*
Probability	0.1	0.2	0.32	0.18	0.12	0.06	0.02

Table 9.8 Combined probabilities

Lead-time usage	*100*	*200*	*300*	*400*	*500*	*600*	*700*	*800*
Probability of usage being greater than X	01.0	0.9	0.7	0.38	0.2	0.08	0.02	0

Continuous and periodic review

The approach we have described is often called the continuous review approach. To make the decision in this way, the stock level of each item must be reviewed continuously and an order placed when the stock level reaches its reorder level. The virtue of this approach is that, although the timing of orders may be irregular (depending on the variation in demand rate), the order size (Q) is constant and can be set at the optimum economic order quantity. But continually checking inventory levels may be time-consuming. An alternative, and simpler, approach, but one that sacrifices the use of a fixed (and therefore possibly optimum) order quantity, is 'periodic review'. Here, rather than ordering at a predetermined reorder level, the periodic approach orders at a fixed and regular time interval. So the stock level of an item could be found, for example, at the end of every month and a replenishment order placed to bring the stock up to a predetermined level. This level is calculated to cover demand between the replenishment order being placed and the following replenishment order arriving. Safety stocks will also need to be calculated, in a similar manner to before, based on the distribution of usage over this period.

Two-bin and three-bin systems

Keeping track of inventory levels is especially important in continuous review approaches to reordering. A simple and obvious method of indicating when the reorder point has been reached is necessary, especially if there are a large number of items to be monitored. The simple two-bin system involves storing the reorder point quantity plus the safety inventory quantity in the second bin and using parts from the first bin. When the first bin empties, it is the signal to order the next reorder quantity. Different 'bins' are not always necessary to operate this type of system. For example, a common practice in retail operations is to store the second 'bin' quantity upside-down behind or under the first 'bin' quantity. Orders are then placed when the upside-down items are reached.

9.5 Diagnostic question: *Is inventory being controlled effectively?*

Even probabilistic models are still simplified compared with the complexity of real stock management. Coping with many thousands of stocked items, supplied by many hundreds of different suppliers, with possibly tens of thousands of individual customers, makes for a complex and dynamic operations task. Controlling such complexity requires an approach that discriminates between different items so that each has a degree of control that is appropriate to its importance. It also requires an information system to keep track of inventories.

Inventory priorities – the ABC system

Some stocked items are more important than others. Some might have a high usage rate, so if they ran out many customers would be disappointed. Others might be of particularly high value, so excessively high inventory levels would be particularly expensive. One common way of discriminating between different stock items is to rank them by their usage value (usage rate multiplied by value). Items with a particularly high usage value are deemed to warrant the most careful control, whereas those with low usage values need not be controlled quite so rigorously. Generally, a relatively small proportion of the total range of items contained in an inventory will account for a large proportion of the total usage value. This phenomenon is known as the Pareto, or 80/20 rule. It is called this because, typically, 80 per cent of an operation's sales are accounted for by only 20 per cent of all stocked item types. (This idea is also used elsewhere in operations managment – see, for example, Chapter 12.) Here the relationship is used to classify items into A, B, or C categories, depending on their usage value:

- Class A items are those 20 per cent or so of high-usage-value items that account for around 80 per cent of the total usage value.
- Class B items are those of medium usage value, usually the next 30 per cent of items, which often account for around 10 per cent of the total usage value.
- Class C items are those low-usage-value items which, although comprising around 50 per cent of the total types of items stocked, probably only account for around 10 per cent of the total usage value of the operation.

> OPERATIONS PRINCIPLE
> ***Different inventory management decision rules are needed for different classes of inventory.***

Although annual usage and value are the two criteria most commonly used to determine a stock classification system, other criteria might also contribute towards the (higher) classification of an item. The consequence of a stock-out might give higher priority to those items that would seriously delay or disrupt operations if they were not in stock. Uncertainty of supply may also give some items priority, as might high obsolescence or deterioration risk.

Worked example

Selectro electrical wholesaler

Table 9.9 shows all the parts stored by Selectro, an electrical wholesaler. The 20 different items stored vary in terms of both their usage per year and cost per item, as shown. However, the wholesaler has ranked the stock items by their usage value per year. The total usage value per year is £5,569,000. From this it is possible to calculate the usage value per year of each item as a percentage of the total usage value, and from that a running cumulative total of the usage value as shown. The wholesaler can then plot the cumulative percentage of all stocked items against the cumulative percentage of their value. So, for example, the part with stock number A/703 is the highest-value part and accounts for 25.14 per cent of the total inventory value. As a part, however, it is only one-twentieth (or 5 per cent) of the total number of items stocked. This item, together with the next-highest-value item (D/012), account for only 10 per cent of the total number of items stocked, yet account for 47.37 per cent of the value of the stock, and so on.

This is shown graphically in Figure 9.11. The first four part numbers (20 per cent of the range) are considered Class A, whose usage will be monitored very closely. The next six part numbers (30 per cent of the range) are to be treated as Class B items, with slightly less effort devoted to their control. All other items are classed as Class C items, whose stocking policy is reviewed only occasionally.

Table 9.9 Warehouse items ranked by usage value

Stock no.	*Usage (items/year)*	*Cost (£/item)*	*Usage value (£000/year)*	*% of total value*	*Cumulative % of total value*
A/703	700	20.00	1,400	25.41	25.14
D/012	450	2.75	1,238	22.23	47.37
A/135	1000	0.90	900	16.16	63.53
C/732	95	8.50	808	14.51	78.04
C/735	520	0.54	281	5.05	83.09
A/500	73	2.30	168	3.02	86.11
D/111	520	0.22	114	2.05	88.16
D/231	170	0.65	111	1.99	90.15
E/781	250	0.34	85	1.53	91.68
A/138	250	0.30	75	1.34	93.02
D/175	400	0.14	56	1.01	94.03
E/001	80	0.63	50	0.89	94.92
C/150	230	0.21	48	0.86	95.78
F/030	400	0.12	48	0.86	96.64
D/703	500	0.09	45	0.81	97.45
D/535	50	0.88	44	0.79	98.24
C/541	70	0.57	40	0.71	98.95
A/260	50	0.64	32	0.57	99.52
B/141	50	0.32	16	0.28	99.80
D/021	20	0.50	10	0.20	100.00
Total			**5,569**	**100.00**	

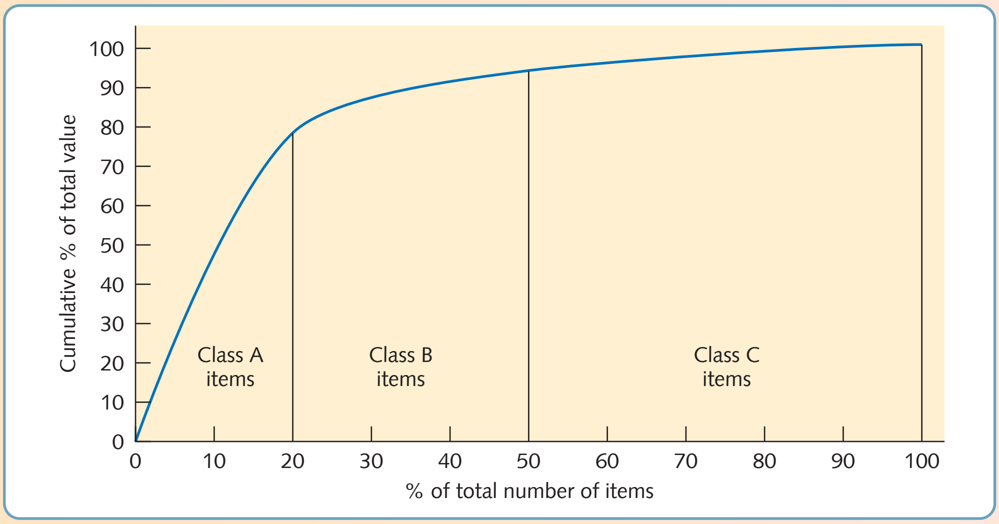


Figure 9.11 **Pareto curve for items in a warehouse**

Inventory information systems

Most inventories of any significant size are managed by information systems. This is especially so since data capture has been made more convenient through the use of bar-code readers, radio frequency identification (RFID) and the point-of-sale recording of sales transactions. Many commercial systems of stock control are available, although they tend to share certain common functions.

Updating stock records – Every time an inventory transaction takes place, the position, status and possibly value of the stock will have changed. This information must be recorded so that operations managers can determine their current inventory status at any time.

Generating orders – Both the how much and the when to order decisions can be made by a stock control system. Originally almost all computer systems calculated order quantities by using the EOQ formulae. Now more sophisticated probabilistic algorithms are used, based on examining the marginal return on investing in stock. The system will hold all the information that goes into the ordering algorithm, but might periodically check to see if demand or order lead times, or any of the other parameters, have changed significantly and recalculate accordingly. The decision on when to order, on the other hand, is a far more routine affair, which computer systems make according to whatever decision rules operations managers have chosen to adopt: either continuous review or periodic review.

Generating inventory reports – Inventory control systems can generate regular reports of stock value, which can help management monitor its inventory control performance. Similarly, customer service performance, such as the number of stock-outs or the number of incomplete orders, can be regularly monitored. Some reports may be generated on an exception basis – that is, the report is generated only if some performance measure deviates from acceptable limits.

Forecasting – Inventory replenishment decisions should ideally be made with a clear understanding of forecast future demand. Inventory control systems usually compare actual demand against forecast and adjust forecasts in the light of actual levels of demand.

Case example

France to ban the dumping of unsold stock[4]

The French government became the first to ban retailers from destroying unsold stock when it announced that the practice will be outlawed by 2021 for most goods and by 2023 for all goods. Brune Poirson, the ecology minister, said that stock worth €800 million went unsold in France every year, of which only €140 millions' worth was given to charities and the rest was destroyed. It had become a widespread practice because occasionally businesses end up with either too little or too much inventory to serve their markets. Too little inventory will result in reduced customer service. But too much can be even more problematic, particularly for some businesses that trade in high-value, 'brand integrity' goods. What does a business do when demand slows and it can't sell any surplus stock without affecting its brand? In 2016, Burberry, the upmarket fashion brand, had to defend its decision to destroy £19 millions' worth of its products that it could not sell through its discount outlet stores. At its annual meeting in London, the company said that it was looking to reduce the amount of wasted stock 'every single season', but also said that destroying surplus stock was a common practice among luxury goods companies. The company's outgoing chief executive said: '*We have a process where we have a sale, then packs go to [a discount] outlet . . . There are some raw materials at the end of that process that we do have to destroy because of intellectual property. It's a common practice but it's something we're enormously mindful of. Every single season we look at how we can reduce, and we have reduced it over the years.*'

Burberry is not alone. When sales of Cartier and Montblanc products slowed sharply, partly because of a crackdown on corruption in China, the overstocking by dealers and an uncertain outlook for growth, the Swiss luxury group Richemont that owns the brands bought back stock from some of its Hong Kong dealers. The watches that were bought back were either reallocated to other regions or, in the case of older models that were no longer selling, were dismantled and recycled. With some luxury goods, the tax rules in some countries actively encourage scrapping surplus stock. For example, if a company makes a bottle of perfume, its cost is a relatively tiny amount (the value comes through advertising and the effect it has on public perception). But the tax loss that the company can claim comes from destroying the product is based on its retail price, not production cost. Of course, there are perfectly legitimate reasons for destroying surplus stock. Any business is responsible for protecting its intellectual property and its brand. However, stock destruction as a means of maintaining 'brand integrity' can backfire. After bags of slashed and cut clothing were found outside one of its New York stores, the clothing retailer H&M had to promise that it would stop destroying new, unworn clothing that it could not sell, and would instead donate the garments to charities.

Helen King/The Image Bank/Getty Images

Common problems with inventory systems

Our description of inventory systems has been based on the assumption that operations:

- have a reasonably accurate idea of costs, such as holding cost or order cost; and
- have accurate information that really does indicate the actual level of stock and sales.

In fact, data inaccuracy often poses one of the most significant problems for inventory managers. This is because most computer-based inventory management systems are based on what is called the perpetual inventory principle. This is the simple idea that stock records are (or should be) automatically updated every time that items are recorded as having been received into an inventory or taken out of the inventory. So,

$$\text{opening stock level} + \text{receipts in} - \text{dispatches out} = \text{new stock level}$$

Any errors in recording these transactions, and/or in handling the physical inventory, can lead to discrepancies between the recorded and actual inventory, and these errors are perpetuated

> OPERATIONS PRINCIPLE
> *The maintenance of data accuracy is vital for the day-to-day effectiveness of inventory management systems.*

until physical stock checks are made (which is usually quite infrequently). In practice there are many opportunities for errors to occur, if only because inventory transactions are numerous. This means that it is surprisingly common for the majority of inventory records to be inaccurate. The underlying causes of errors include:

- keying errors – entering the wrong product code;
- quantity errors – a miscount of items put into or taken from stock;
- damaged or deteriorated inventory not recorded as such, or not correctly deleted from the records when it is destroyed;
- the wrong items being taken out of stock, but the records not being corrected when they are returned to stock;
- delays between the transactions being made and the records being updated;
- items stolen from inventory (common in retail environments, but also not unusual in industrial and commercial inventories).

Critical commentary

- The approach to determining order quantity that involves optimising costs of holding stock against costs of ordering stock, typified by the EOQ and EBQ models, has always been subject to criticisms. Originally these concerned the validity of some of the assumptions of the model; more recently they have involved the underlying rationale of the approach itself. Criticisms are that the assumptions included in the EOQ models are simplistic, that the real costs of stock in operations are not as assumed in EOQ models, and that cost minimisation is not an appropriate objective for inventory management.
- The last criticism is particularly significant. Many organisations (such as supermarkets and wholesalers) make most of their revenue and profits simply by holding and supplying inventory. Because their main investments are in the inventory, it is critical that they make a good return on this capital, by ensuring that it has the highest possible 'stock turn' and/or gross profit margin. Alternatively, they may also be concerned to maximise the use of space by seeking to maximise the profit earned per square metre. The EOQ model does not address these objectives. Similarly, for products that deteriorate or go out of fashion, the EOQ model can result in excess inventory of slower-moving items. In fact, the EOQ model is rarely used in such organisations, and there is more likely to be a system of periodic review for regular ordering of replenishment inventory. For example, a typical builders' supply merchant might carry around 50,000 different items of stock (SKUs). However, most of these cluster into larger families of items such as paints, sanitary ware or metal fixings. Single orders are placed at regular intervals for all the required replenishments in the supplier's range, and these are then delivered together at one time. If deliveries are made weekly then, on average, the individual item order quantities will be for only one week's usage. Less popular items, or ones with erratic demand patterns, can be ordered individually at the same time, or (when urgent) can be delivered the next day by carrier.
- The ABC approach to inventory classification is regarded by some as misleading. Many professional inventory managers point out that it is the slow-moving (C-category) items that often pose the greatest challenge in inventory management. Often these slow-moving items, although accounting for only 20 per cent of sales, require a large part (typically between one half and two thirds) of the total investment in stock. This is why slow-moving items are a real problem. Moreover, if errors in forecasting or ordering result in excess stock in 'A-class' fast-moving items, it is relatively unimportant in the sense that excess stock can be sold quickly. However, excess stock in slow-moving C items will be there a long time. According to some inventory managers, it is the A items that can be left to look after themselves; it is the B and, even more so, the C items that need controlling.

SUMMARY CHECKLIST

- ☐ Have all inventories been itemised and costed?
- ☐ Have all the costs and negative effects of inventory been assessed?
- ☐ What proportion of inventory is there:
  - as an insurance against uncertainty?
  - to counteract a lack of flexibility?
  - to allow operations to take advantage of short-term opportunities?
  - to anticipate future demand?
  - to reduce overall costs?
  - because it can increase in value?
  - because it is in the processing pipeline?
- ☐ Have methods of reducing inventory in these categories been explored?
- ☐ Have cost-minimisation methods been used to determine order quantity?
- ☐ Do these use a probabilistic estimate of demand?
- ☐ Have the relative merits of continuous and periodic inventory review been assessed?
- ☐ Are probabilistic estimates of demand and lead time used to determine safety stock levels?
- ☐ Are items controlled according to their usage value?
- ☐ Does the inventory information system integrate all inventory decisions?

Case study

supplies4medics.com

Founded more than 20 years ago, supplies4medics.com has become one of Europe's most successful direct-mail suppliers of medical hardware and consumables to hospitals, doctors' and dentists' surgeries, clinics, nursing homes and other medical-related organisations. Its physical and online catalogues list just over 4,000 items, categorised by broad applications such as 'hygiene consumables' and 'surgeons' instruments'. Quoting their website:

> *'We are the pan-European distributors of wholesale medical and safety supplies ... We aim to carry everything you might ever need; from nurses' scrubs to medical kits, consumables for operations, first-aid kits, safety products, chemicals, fire-fighting equipment, nurse and physicians' supplies, etc. Everything is at affordable prices – and backed by our very superior customer service and support –* **supplies4medics** *is your ideal source for all medical supplies. Orders are normally dispatched same-day, via our European distribution partner, the Brussels hub of DHL. You should therefore receive your complete order within one week, but you can request next-day delivery if required, for a small extra charge. You can order our printed catalogue through the link at the bottom of this page, or shop on our easy-to-use online store.'*

Last year turnover grew by over 25 per cent to about €120 million, a cause for considerable satisfaction in the company. However, profit growth was less spectacular; and market research suggested that customer satisfaction, although generally good, was slowly declining. Most worryingly, inventory levels had grown faster than sales revenue, in percentage terms. This was putting a strain on cash flow, requiring the company to borrow more cash to fund the rapid growth planned for the next year. Inventory holding was estimated to be costing around 15 per cent per annum, taking account of the cost of borrowing, insurance and all warehousing overheads.

Pierre Lamouche, the Head of Operations, summarised the situation faced by his department:

> *'As a matter of urgency, we are reviewing our purchasing and inventory management systems! Most of our existing reorder levels (ROL) and reorder quantities (ROQ) were set several years ago, and have never been recalculated. Our focus has been on rapid growth through the introduction of new product lines. For more recently introduced items, the ROQs were based only on forecast sales, which can actually be quite misleading. We estimate that it costs us, on average, €50 to place and administer every purchase order, since most suppliers are still not able to take orders over the internet or by EDI. In the meantime, sales of some products have grown fast, while others have declined. Our average inventory (stock) cover is about 10 weeks, but, amazingly, we still run out of critical items! In fact, on average, we are currently out of stock of about 500 SKUs (stock keeping units) at any time. As you can imagine, our service level is not always satisfactory with this situation. We really need help to conduct a review of our system, so have employed a mature intern from the local business school. He has first asked my team to provide information on a random, representative sample of 20 items from the full catalogue range, which is copied below.'* (See Table 9.10)

Charlie Neuman/San Diego Union-Tribune/Zuma Press Inc/Alamy Stock Photo

Questions

1. Prepare a spreadsheet-based ABC analysis of usage value.
2. Calculate the inventory weeks for each item, for each classification, and for all the items in total. Does this suggest that the Head of Operations' estimate of inventory weeks is correct? If so, what is your estimate of the overall inventory at the end of the base year, and how much might that have increased during the year?
3. Based on the sample, analyse the underlying causes of the availability problem described in the text.
4. Calculate the EOQs for the A items.
5. What recommendations would you make to the company?

Table 9.10 Representative sample of 20 catalogue items

Sample number	*Catalogue reference number**	*Sales unit description ***	*Sales unit cost (€)*	*Last 12 months' sales (units)*	*Inventory as at last year-end (units)*	*Reorder quantity (units)*
1	11036	Disposable aprons (10pk)	2.40	100	0	10
2	11456	Ear-loop masks (box)	3.60	6,000	1,200	1,000
3	11563	Drill type 164	1.10	220	420	250
4	12054	Incontinence pads large	3.50	35,400	8,500	1,0000
5	12372	150ml syringe	11.30	430	120	100
6	12774	Rectal speculum 3-prong	17.40	65	20	20
7	12979	Pocket organiser blue	7.00	120	160	500
8	13063	Oxygen trauma kit	187.00	40	2	10
9	13236	Zinc-oxide tape	1.50	1,260	0	50
10	13454	Dual-head stethoscope	6.25	10	16	25
11	13597	Disp. latex catheter	0.60	3,560	12	20
12	13999	Roll-up wheelchair ramp	152.50	12	44	50
13	14068	WashClene tube	1.40	22,500	10,500	8,000
14	14242	Cervical collar	12.00	140	24	20
15	14310	Head wedge	89.00	44	2	10
16	14405	Three-wheel scooter	755.00	14	5	5
17	14456	Neonatal trach. tube	80.40	268	6	100
18	14675	Mouldable strip paste	10.20	1,250	172	100
19	14854	Sequential comp. pump	430.00	430	40	50
20	24943	Toilet safety frame	25.60	560	18	20

*Reference numbers are allocated sequentially as new items are added to the catalogue.
**All quantities are in sales units (e.g. item, box, case, pack).

Applying the principles

Some of these exercises can be answered by reading the chapter. Others will require some general knowledge of business activity and some might require an element of investigation. **Model answers for the first two questions are available on the student companion website.** *For model answers to the other questions in this section, please ask your tutor/lecturer.*

1. A supplier makes monthly shipments to 'House & Garden Stores', in average lot sizes of 200 coffee tables. The average demand for these items is 50 tables per week, and the lead time from the supplier three weeks. 'House & Garden Stores' must pay for inventory from the moment the supplier ships the products. If they are willing to increase their lot size to 300 units, the supplier will offer a lead time of one week. What will be the effect on cycle and pipeline inventories?

2. A local shop has a relatively stable demand for tins of sweetcorn throughout the year, with an annual total of 1,400 tins. The cost of placing an order is estimated at £15 and the annual cost of holding inventory is estimated at 25 per cent of the product's value. The company purchases tins for 20p. How much should the shop order at a time, and what is the total cost of the plan?

3. A fruit-canning plant has a single line for three different fruit types. Demand for each type of tin is reasonably constant at 50,000 per month (a month has 160 production hours). The tinning process rate is 1,200 per hour, but it takes two hours to clean and reset between different runs. The cost of these changeovers (C_o) is calculated at £250 per hour. Stock-holding is calculated at $0.1 per tin per month. How big should the batch size be?

4. '*Our suppliers often offer better prices if we are willing to buy in larger quantities. This creates a pressure on us to hold higher levels of stock. Therefore, to find the best quantity to order we must compare the advantages of lower prices for purchases and fewer orders with the disadvantages of increased holding costs. This means that calculating total annual inventory-related costs should now not only include holding costs and ordering costs, but also the cost of purchased items themselves*.' (Manager, Tufton Bufton Port Importers, Inc.) One supplier to Tufton Bufton Port Importers, Inc. (TBPI) has introduced quantity discounts to encourage larger order quantities. The discounts are as follows:

Order quantity	*Price per bottle*
0–100	€15.00
101–250	€13.50
250+	€11.00

   TBPI estimates that its annual demand for this particular wine is 1,500 bottles, its ordering costs are €30 per order, and its annual holding costs are 20 per cent of the bottle's price.

   a) How should TBPI go about deciding how many to order?
   b) How many should they order?

5. Most countries have blood collection and distribution services that collect from donors, process the blood by either breaking it down into its constituent parts or keeping it whole, and transporting the blood from collection centres to hospitals in response to both routine and emergency requests.

   a) What are the factors that constitute inventory-holding costs, order costs and stock-out costs in such a blood service?
   b) What makes this particular example of inventory planning and control so complex?
   c) How might the efficiency with which a blood service controls its inventory affect its ability to collect blood?

Notes on chapter

1 The information on which this example is based is taken from: Aldrick, P, (2020) 'Coronavirus: PPE stocks were in chaos . . . then army got a grip', *The Times*, 11 June; Foster, P. and Neville, S. (2020) 'How poor planning left the UK without enough PPE', *Financial Times*, 1 May; Brit, H. (2020) 'What is safety stock and how can businesses use it to ensure continuity?', Thomasnet.com, 8 April, https://www.thomasnet.com/insights/what-is-safety-stock/ [accessed 23 September 2020]; Anderson, H. (2020) 'COVID-19: Preparing your supply chain in times of crisis', publicissapient.com, 8 April, https://www.publicissapient.com/insights/coronavirus_and_managing_the_supply_chain_amid_a_crisis [accessed 23 September 2020].

2 The information on which this example is based is taken from: Ulanoff, L. (2014) 'Amazon knows what you want before you buy it', Mashable, https://mashable.com/2014/01/21/amazon-anticipatory-shipping-patent/#Ryy4twKmRiqb [accessed 23 September 2020]; Duke, S. (2014) 'He knows what you want – before you even want it', *Sunday Times*, 2 February; Ahmed, M. (2014) 'Amazon will know what you want before you do', *The Times*, 27 January; Bernard, Z. (2018) 'Amazon is spending more and more on shipping out your orders', *Business Insider*.

3 The information on which this example is based is taken from: Economist (2015) 'Croissantonomics: Lessons in managing supply and demand for perishable products', *Economist* print edition, 29 August.

4 The information on which this example is based is taken from: Sage, A. (2019) 'France to ban luxury brands from dumping unsold stock', *The Times*, 24 September; Leroux, M. (2016) 'Burberry boss defends stock destruction', *The Times*, 15 July; Atkins, R. (2016) 'Richemont buys back and destroys stock as sales fall', *Financial Times*, 20 May; Dwyer, J. (2010) 'A clothing clearance where more than just the prices have been slashed', *New York Times*, 5 January.

Taking it further

Bragg, S.M. ***(2018) Inventory Management, 3rd edition, AccountingTools, Inc.*** *A financial approach to the subject.*

Emmett, S. and Granville, D. ***(2007) Excellence in Inventory Management: How to minimise costs and maximise service, Cambridge Academic.*** *Practical and thorough examination of professional inventory management.*

Muller, M. ***(2011) Essentials of Inventory Management, 2nd edition, Amacom.*** *A good introduction.*

Relph, G. and Milner, C. ***(2019) The Inventory Toolkit: Business systems solutions, 2nd edition, Kogan Page.*** *Cuts through the complexity of inventory theory.*

Wild, T. ***(2017) Best Practice in Inventory Management, 3rd edition, Routledge.*** *Long-standing and authoritative text.*

10 Resource planning and control

Introduction

After the nature of an operation's resources has been at least approximately determined, the activities that create its services and products must be managed on an ongoing basis. This is the activity of resource planning and control. It is concerned with what activities happen in operations, when they happen, where they happen and what resources are going to be involved. It ensures that materials or information or customers flow smoothly through processes, operations and supply networks, and that value-adding resources are managed efficiently and to avoid unnecessary delay. Resource planning and control is a subject with many technical issues. We cover the best known of these (materials requirements planning, or MRP in the supplement to this chapter. Figure 10.1 shows the position of the ideas described in this chapter within the general model of operations management.

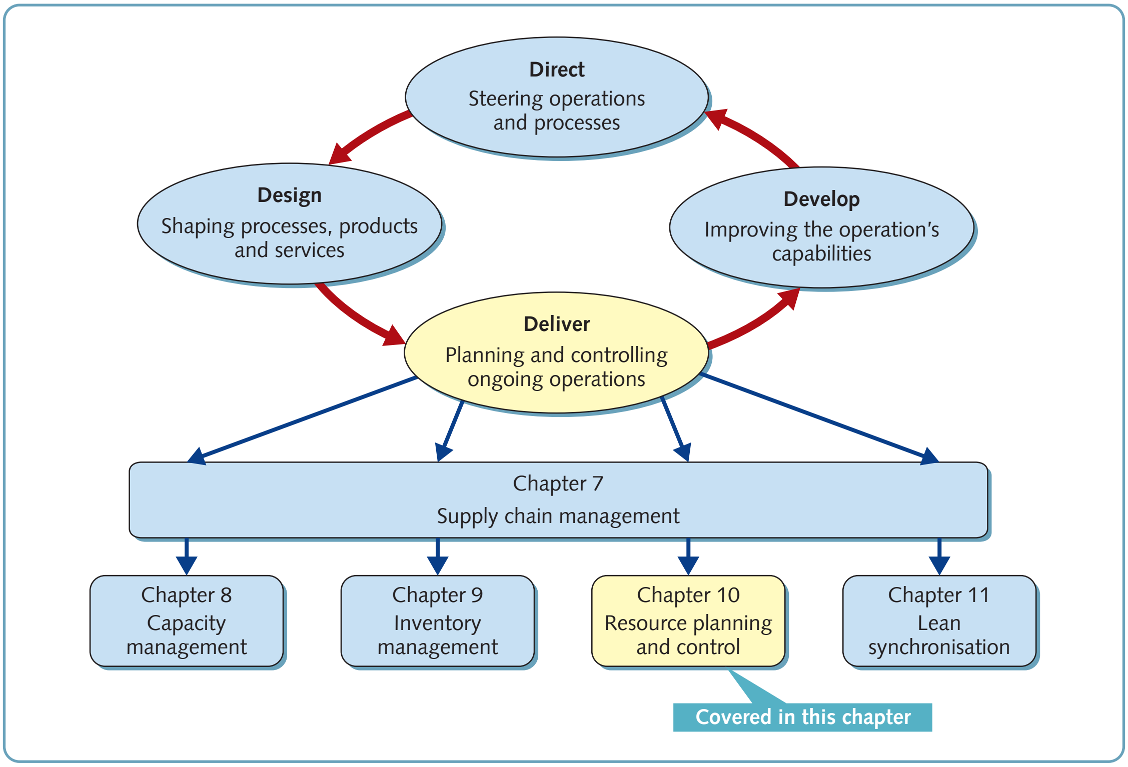


Figure 10.1 Resource planning and control is concerned with managing the ongoing allocation of resources and activities to ensure that the operation's processes are both efficient and reflect customer demand for products and services

EXECUTIVE SUMMARY

10.1 Does resource planning and control have all the right elements?

10.2 Is resource planning and control information integrated?

10.3 Are core planning and control activities effective?

10.1 Does resource planning and control have all the right elements?

Resource planning and control is concerned with managing the ongoing allocation of resources and activities to ensure that the operation's processes are both efficient and reflect customer demand for products and services. In practice, planning (deciding what is intended to happen) and control (coping when things do not happen as intended) overlap to such an extent that they are usually treated together.

Although planning and control systems differ, they tend to have a number of common elements. These are: a customer interface that forms a two-way information link between the operation's activities and its customers; a supply interface that does the same thing for the operation's suppliers; a set of overlapping 'core' mechanisms that perform basic tasks such as loading, sequencing, scheduling, and monitoring and control; and a decision mechanism involving both operations staff and information systems that makes or confirms planning and control decisions. It is important that all these elements are effective in their own right and work well together.

10.2 Is resource planning and control information integrated?

Resource planning and control involves vast amounts of information. Unless all relevant information is integrated it is difficult to make informed planning and control decisions. The most common method of doing this is through the use of integrated 'enterprise resource planning' (ERP) systems. These are information systems that have grown out of the more specialised and detailed material requirements planning (MRP) systems that have been common in the manufacturing sector for many years. MRP is treated in the supplement to this chapter. Investment in ERP systems often involves large amounts of capital and staff time. It also may mean a significant overhaul of the way the business organises itself. Not all investments in ERP have proved successful.

10.3 Are core planning and control activities effective?

Unless the resource planning and control system makes appropriate decisions at a detailed level, it cannot be effective. These detailed decisions fall into four overlapping categories. Loading is the activity of allocating work to individual processes or stages in the operation. Sequencing is the activity of deciding the order or priority in which a number of jobs will be processed. Scheduling is the activity of producing a detailed timetable showing when activities should start and end. Monitoring and control is the activity of detecting any deviation from what has been planned and acting to cope and replan as necessary. The theory of constraints (TOC) is a useful concept in resource planning and control that emphasises the role of bottleneck stages or processes in planning and control.

10.1 Diagnostic question: *Does resource planning and control have all the right elements?*

Resource planning and control is concerned with managing the ongoing allocation of resources and activities to ensure that the operation's processes are both efficient and reflect customer demand for products and services. Planning and control activities are distinct but often overlap. Formally, planning determines what is *intended* to happen at some time in the future, while control is the process of *coping* when things do not happen as intended. Control makes the adjustments that help the operation to achieve the objectives that the plan has set, even when the assumptions on which the plan was based do not hold true.

Case example

Operations control at Air France[1]

'In many ways a major airline can be viewed as one large planning problem which is usually approached as many independent, smaller (but still difficult) planning problems. The list of things which need planning seems endless: crews, reservation agents, luggage, flights, through trips, maintenance, gates, inventory, equipment purchases. Each planning problem has its own considerations, its own complexities, its own set of time horizons, its own objectives, but all are interrelated.' (Rikard Monet, Air France)

Air France has 80 flight planners working 24-hour shifts in their flight planning office at Roissy, Charles de Gaulle. Their job is to establish the optimum flight routes, anticipate any problems such as weather changes, and minimise fuel consumption. Overall, the goals of the flight planning activity are first, and most importantly, safety, followed by economy and passenger comfort. Increasingly powerful computer programs process the mountain of data necessary to plan the flights, but in the end many decisions still rely on human judgement. Even the most sophisticated expert systems only serve as support for the flight planners. Planning Air France's schedule is a massive job that includes the following:

- *Frequency* – for each airport, how many separate services should the airline provide?
- *Fleet assignment* – which type of plane should be used on each leg of a flight?
- *Banks* – at any airline hub where passengers arrive and may transfer to other flights to continue their journey, airlines like to organise flights into 'banks' of several planes, which arrive close together, pause to let passengers change planes and all depart close together.
- *Block times* – a block time is the elapsed time between a plane leaving the departure gate at an airport and arriving at its gate in the arrival airport. The longer the allowed block time, the more likely a plane will keep to schedule even if it suffers minor delays, but fewer flights can be scheduled.
- *Planned maintenance* – any schedule must allow time for planes to be worked on at a maintenance base.
- *Crew planning* – pilot and cabin crew must be scheduled to allocate pilots to fly planes on which they are licensed and to keep within the maximum 'on duty' allowances.
- *Gate plotting* – if many planes are on the ground at the same time there may be problems in loading and unloading them simultaneously.
- *Recovery* – many things can cause deviations from any plan in the airline industry. Allowances must be built in that allow for recovery.

For flights within and between Air France's 12 geographic zones, the planners construct a flight plan that will form the basis of the actual flight only a few hours later. All planning documents need to be ready for the flight crew who arrive two hours before the scheduled departure time. Being responsible for passenger safety and comfort, the captain always has the final say and, when satisfied, co-signs the flight plan together with the planning officer.

EQRoy/Shutterstock

The planning and control system described in the case example 'Operations control at Air France' has a number of elements common to all effective systems. First, there is some kind of acknowledgement that there should be an effective *customer interface* that translates the needs of customers into their implications for the operation. This involves setting the timetable of flights (frequency, timing, etc.) and the interfaces between flights (banks). Their planning and control system also has a *supply interface* that translates the operation's plans in terms of the supply of parts, or fuel, ground services, crew availability, etc. At the heart of the system is a set of *core mechanics* that loads capacity, prioritises jobs, schedules work, and monitors and controls the operation. The job of this decision-making is to reconcile the needs of the customers and the operation's resources in some way. For Air France, customer comfort and safety are paramount. The system also attempts some *information integration* that involves both computer-assisted information handling and the skills and experience of planning and control staff. This chapter covers these elements, common to all planning and control.

What elements should all resource planning and control have?

Figure 10.2 illustrates the elements that should be present in all planning and control systems. In more sophisticated systems they may even be extended to include the integration of this core operations resource planning and control task with other functional areas of the firm such as finance, marketing and personnel. We deal with this cross-functional perspective when we discuss enterprise resource planning (ERP) later.

How does the system interface with customers?

The part of the resource planning and control system that manages the way customers interact with the business on a day-to-day basis is called the 'customer interface' or sometimes 'demand management'. This is a set of activities that interface with both individual customers and the market more broadly. Depending on the business, these activities may include customer negotiation, order entry, demand forecasting, order promising, updating customers, keeping customer histories, post-delivery customer service and physical distribution.

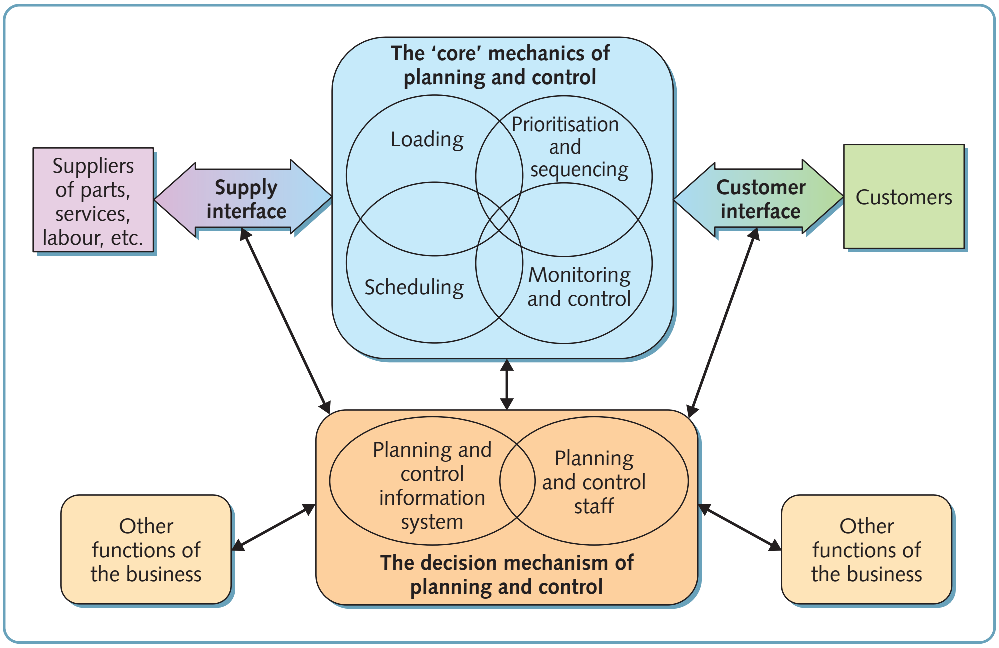


Figure 10.2 **The key elements of a resource planning and control system**

Customer interface defines the customer experience

The customer interface is important because it defines the nature of the customer experience. It is the public face of the operation (the 'line of visibility' as it was called in Chapter 6). Therefore, it needs to be managed like any other 'customer processing' process, where the quality of the service, as the customer sees it, is defined by the gap between customers' expectations and their perceptions of the service they receive. Figure 10.3 illustrates a typical customer experience of interacting with a planning and control customer interface. The experience itself will start before any customer contact is initiated. Customer expectations will have been influenced by the way the business presents itself through promotional activities, the ease with which channels of communication can be used (for example, design of the website), and so on. The question is 'does the communication channel give any indication of the kind of service response (for example, how long we will have to wait) that the customer can expect'. At the first point of contact, when an individual customer requests services or products, their request must be understood, delivery possibly negotiated, and a delivery promise made. Prior to the delivery of the service or product, the customer may or may not change their mind, which in turn may or may not involve renegotiation of delivery promises. Similarly, customers may require or value feedback on the progress of their request. At the point of delivery, not only are the products and services handed over to the customer, there may also be an opportunity to explain the nature of the delivery and gauge customers' reactions. Following the completion of the delivery there may also be some post-delivery action, such as a phone call to confirm that all is well.

> OPERATIONS PRINCIPLE
> ***Customers' perceptions of an operation will partially be shaped by the customer interface of its planning and control system.***

As is usual with such customer experiences, the managing of customer expectations is particularly important in the early stages of the experience. For example, if there is a possibility that a delivery may be late (perhaps because of the nature of the service being requested), then that possibility is established as an element in the customer's expectations. As the experience continues, various interactions with the customer interface service build up customer perceptions of the level of support and care exhibited by the operation. The right-hand side of Figure 10.3

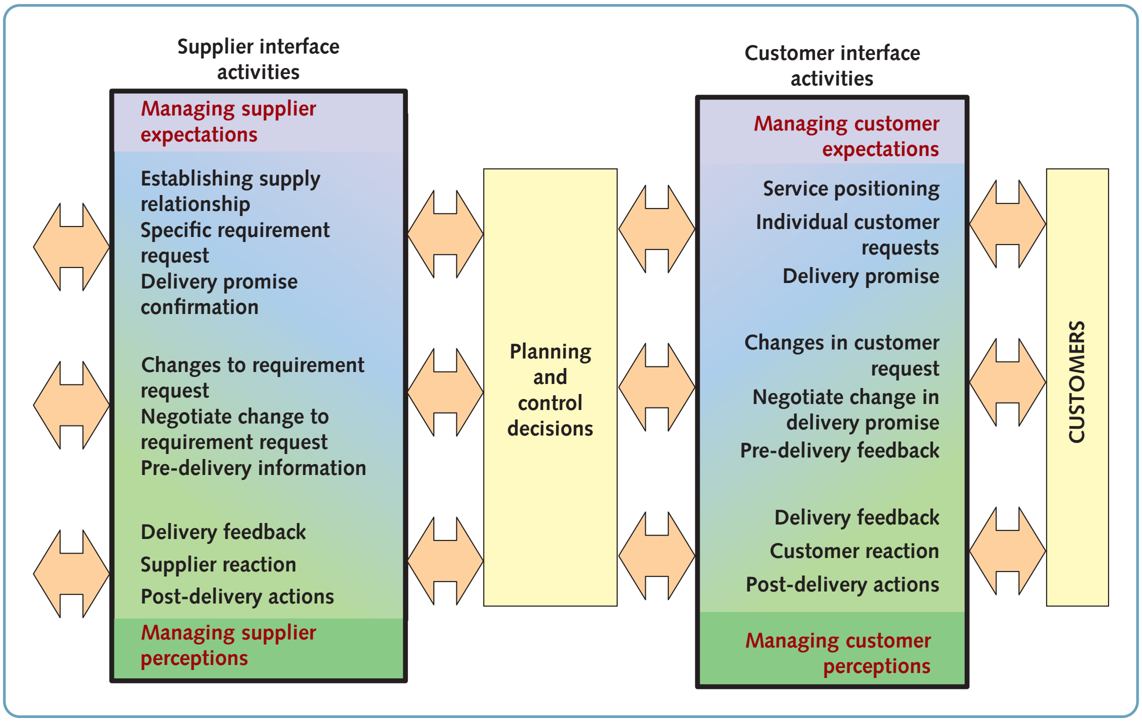


Figure 10.3 The customer and supplier interfaces as a 'customer experience'

shows a simplified sequence of events in the management of a typical customer–operation interaction that the customer interface must facilitate.

The customer interface should reflect the operation's objectives

In managing a customer's experience, the customer interface element of the planning and control system is, in effect, operationalising the business's operations objectives. It may have to prioritise one type of customer over another. It may have to encourage some types of customers to transact business more than other (possibly less profitable) types of customers. It will almost certainly have to trade off elements of customer service against the efficiency and utilisation of the operations resources. No matter how sophisticated the customer interface technology, or how skilled the customer interface staff, this part of the planning and control system cannot operate effectively without clear priorities derived from the operation's strategic objectives.

The customer interface acts as a trigger function

Acceptance of an order should prompt the customer interface to trigger the operation's processes. Exactly what is triggered will depend on the nature of the business. For example, some building and construction companies, because they are willing to build almost any kind of construction, will keep relatively few of their own resources within the business, but rather hire them in when the nature of the job becomes evident. This is a 'resource-to-order' operation where the customer interface triggers the task of hiring in the relevant equipment (and possibly labour) and purchasing the appropriate materials. If the construction company confines itself to a narrower range of construction tasks, thereby making the nature of demand slightly more predictable, it will be likely to have its own equipment and labour permanently within the operation. Here, accepting a job would only need to trigger the purchase of the materials to be used in the construction, and the business is a 'produce-to-order' operation. Some construction companies will construct pre-designed standard houses or apartments ahead of any firm demand for them. If demand is high, customers may place requests for houses before they are started or during their construction. In this case, the customer will form a backlog of demand and must wait. However, the company is also taking the risk of holding a stock of unsold houses. Operations of this type 'produce ahead of order'.

How does the system interface with suppliers?

The supplier interface provides the link between the activities of the operation itself and those of its suppliers. The timing and level of activities within the operation or process will have implications for the supply of products and services to the operation. Suppliers need to be informed so that they can make products and services available when needed. In effect this is the mirror image of the customer interface. As such, the supplier interface is concerned with managing the supplier experience to ensure appropriate supply. Because the customer is not directly involved in this does not make it any less important. Ultimately, customer satisfaction will be influenced by supply effectiveness because that in turn influences delivery to customers. Using the expectations–perception gap to judge the quality of the supplier interface function may at first seem strange. After all, suppliers are not customers as such. Yet, it is important to be a 'quality customer' to suppliers because this increases the chances of receiving high-quality service from them. This means that suppliers fully understand the operation's expectations because they have been made clear and unambiguous.

> **OPERATIONS PRINCIPLE**
> ***An operation's planning and control system can enhance or inhibit the ability of its suppliers to support delivery effectiveness.***

The supplier interface has both a long- and short-term function. It must be able to cope with different types of long-term supplier relationship, and also handle individual transactions with suppliers. To do the former, it must understand the requirements of all the processes within the operation, and also the capabilities of the suppliers (in large operations, there could be thousands of suppliers). The left-hand side of Figure 10.3 shows a simplified sequence of events in the management of a typical supplier–operation interaction, which the supplier interface must facilitate. When the planning and control activity requests supply, the supplier interface must have identified potential suppliers and might also be able to suggest alternative materials or services if necessary. Formal requests for quotations may be sent to potential suppliers if no supply agreement exists. These requests might be sent to several suppliers or a smaller group, who may be 'preferred' suppliers. Just as it was important to manage customer expectations, it is important to manage supplier expectations, often prior to any formal supply of products or services. This issue was discussed in Chapter 7 as supplier development. To handle individual transactions, the supplier interface will need to issue formal purchase orders. These may be stand-alone documents or, more likely, electronic orders. Whatever the mechanisms, it is an important activity because it often forms the legal basis of the contractual relationship between the operation and its supplier. Delivery promises will need to be formally confirmed. While waiting for delivery, it may be necessary to negotiate changes in supply and track progress to get early warning of potential changes to delivery. Also, supplier delivery performance needs to be established and communicated with follow-up as necessary.

How does the system perform basic planning and control calculations?

Resource planning and control requires the reconciliation of supply and demand in terms of the level and timing of activities within an operation or process. To do this, four overlapping activities are performed. These are loading, sequencing, scheduling, and monitoring and control. However, some caution is needed when using these terms. Different organisations may use them in different ways, and even texts in this area may adopt different definitions. Although these four activities are very closely interrelated, they do address different aspects of the resource planning and control task (see Figure 10.4). Loading allocates tasks to resources in order to assess *what* level of activity will be expected of each part of the operation. Scheduling is more concerned with *when* the operation or process will do things. Sequencing is a more detailed set of decisions that determines *in what order* jobs pass through processes. Monitoring and control involves checking if *activities are going to plan* by observing what is actually happing in practice, and making adjustments as necessary. This part of the planning and control

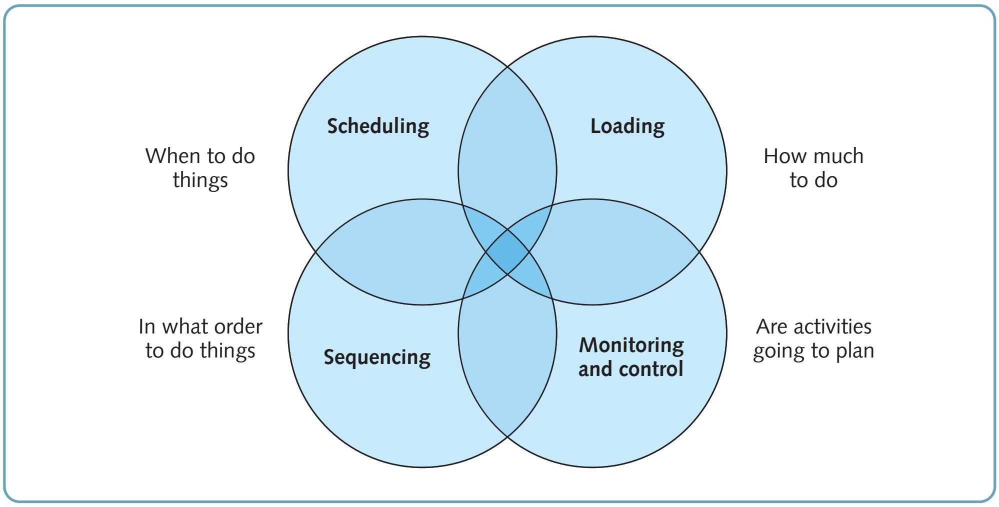


Figure 10.4 The four 'core mechanisms' of planning and control

> **OPERATIONS PRINCIPLE**
> ***Resource planning and control comprises the four core overlapping activities of loading, sequencing, scheduling, and monitoring and control.***

system can be regarded as the engine room of the whole system in so much as it calculates the consequences of planning and control decisions. Without understanding how these basic mechanisms work, it is difficult to understand how any operations are being planned and controlled. Because of their importance, we address the four core interrelated activities later in the chapter.

Does the system integrate human with 'automated' decision-making?

Although computer-based resource planning and control systems are now widespread in many industries, much of the decision-making is still carried out partially by people. This is always likely to be the case because some elements of the task, such as negotiating with customers and suppliers, are difficult to automate. Yet the benefits of computer-aided decision-making are difficult to ignore. Unlike humans, computer-based planning and control can cope with immense complexity, both in terms of being able to model the interrelationship between decisions and in terms of being able to store large quantities of information. However, humans are generally better at many of the 'soft' qualitative tasks that can be important in planning and control. In particular, humans are good at:

- *Flexibility, adaptability and learning.* Humans can cope with ambiguous, incomplete, inconsistent and redundant goals and constraints. In particular, they can deal with the fact that planning and control objectives and constraints may not be stable for longer than a few hours at a time.
- *Communication and negotiation.* Humans are able to understand and sometimes influence the variability inherent in an operation. They can influence job priorities and sometimes processing times. They can negotiate between internal processes and communicate with customers and suppliers in a way that could minimise misunderstandings.
- *Intuition.* Humans can fill in the blanks of missing information that is required to plan and control. They can accumulate the tacit knowledge about what is, and what may be, really happening with the operation's processes.

These strengths of human decision-making versus computer decision-making provide a clue as to what should be the appropriate degree of automation built into decision-making in this area. When planning and controlling stable and relatively straightforward processes that are well understood, decision-making can be automated to a greater degree than for processes that are complex, unstable and poorly understood.

10.2 Diagnostic question: *Is resource planning and control information integrated?*

One of the most important issues in resource planning and control is managing the sometimes vast amounts of information generated, not just from the operations function but from almost every other function of the business. Unless all relevant information is brought together and integrated, it is difficult to make informed planning and control decisions. This is what enterprise resource planning (ERP) is about. It has been defined as '*a complete enterprise-wide business solution. The ERP system consists of software support modules such as: marketing and sales, field service, product design and development, production and inventory control, procurement, distribution, industrial facilities management, process design and development, manufacturing, quality, human resources, finance and accounting, and information services. Integration between the modules is stressed without the duplication of information.*'[2]

> **OPERATIONS PRINCIPLE**
> ***Planning and control systems should integrate information from all relevant organisational functions.***

The origins of ERP

Enterprise resource planning has spawned a huge industry devoted to developing the computer systems needed to drive it. The (now) large companies that have grown almost exclusively on the basis of providing ERP systems include SAP, Oracle and Baan. Yet ERP is just one of the latest (but most important) stages in a development that started with materials requirements planning (MRP), an approach that became popular during the 1970s, although the planning and control logic that underlies it had been known for some time. It is a method (simple in principle but complex in execution) of translating a statement of required output into a plan for all the activities that must take place to achieve the required output. What popularised MRP was the availability of computer power to drive the basic planning and control mathematics in a fast, efficient and, most importantly, flexible manner. MRP is treated in the supplement to this chapter. Manufacturing resource planning II (MRP II) expanded out of MRP during the 1980s. This extended concept has been described as a game plan for planning and monitoring all the resources of a manufacturing company: manufacturing, marketing, finance and engineering. Again, it was a technology innovation that allowed the development. Local area networks (LANs) together with increasingly powerful desktop computers allowed a much higher degree of processing power and communication between different parts of a business.

The strength of MRP, and MRP II, lay always in the fact that it could explore the *consequences* of any changes to what an operation was required to do. So, if demand changed, the MRP system would calculate all the 'knock-on' effects, and issue instructions accordingly. The same principle applies to ERP, but on a much wider basis. ERP systems allow decisions and databases from all parts of the organisation to be integrated so that the consequences of decisions in one part of the organisation are reflected in the planning and control systems of the rest of the organisation (see Figure 10.5).

ERP changes the way companies do business

Arguably the most significant issue in many companies' decision to buy an off-the-shelf ERP system is that of its compatibility with the company's current business processes and practices.

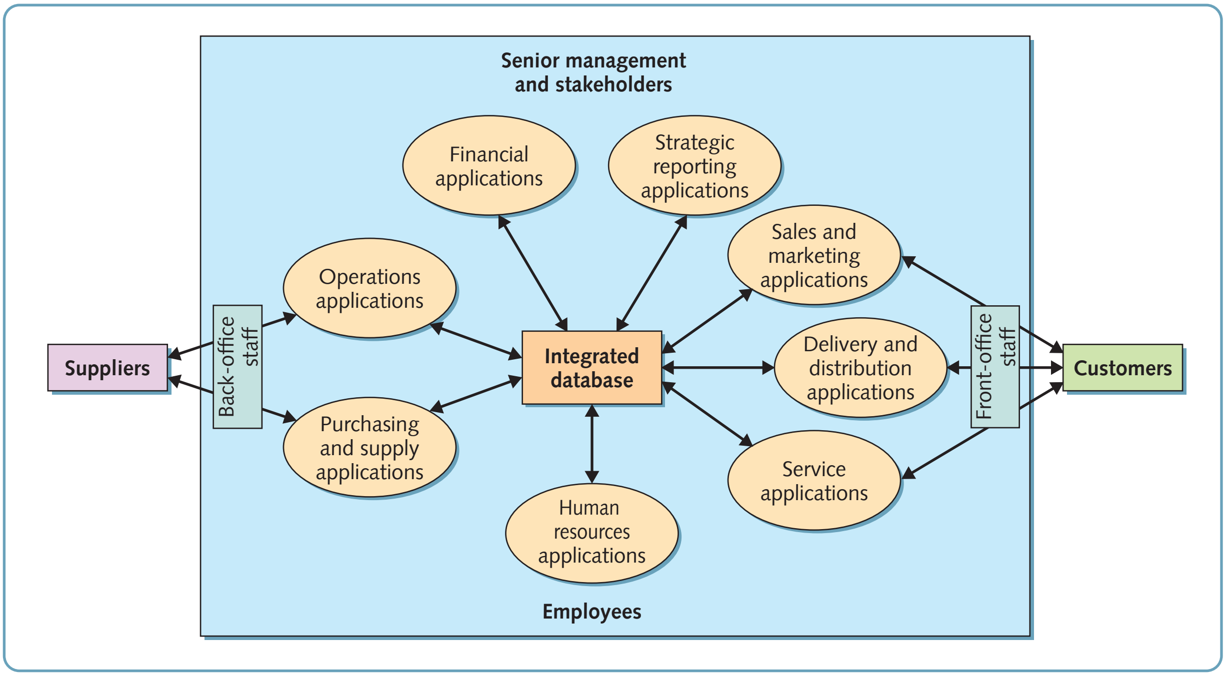


Figure 10.5 **ERP integrates information from all parts of the organisation**

OPERATIONS PRINCIPLE
ERP systems are only fully effective if the way a business organises its processes is aligned with the underlying assumptions of its ERP system.

Experience of ERP installation suggests that it is extremely important to make sure that the company's current way of doing business will fit (or can be changed to fit) with a standard ERP package. One of the most common reasons for not installing ERP is incompatibility between the assumptions in the software and the operating practice of core business processes. If a business's current processes do not fit, it can either change its processes to fit the ERP package, or modify the software within the ERP package to fit its processes, but both of these options involve costs and risks. Changing business practices that are working well will involve reorganisation costs, as well introducing the potential for errors to creep into the processes. Adapting the software will both slow down the project and introduce potentially dangerous software 'bugs' into the system. It will also make it difficult to upgrade the software at a later date.

ERP installation can be particularly expensive. Attempting to get new systems and databases to talk to old (sometimes called *legacy*) systems can be very problematic. Not surprisingly, many companies choose to replace most, if not all, of their existing systems simultaneously. New common systems and relational databases help to ensure the smooth transfer of data between different parts of the organisation.

In addition to the integration of systems, ERP usually includes other features, which make it a powerful planning and control tool. In particular, the availability of web-based and cloud-based ERP services (see later) have made ERP's power available to a wide range of operations of all different sizes.

The benefits of ERP

ERP is generally seen as having the potential to significantly improve the performance of many companies in many different sectors. This is partly because of the very much enhanced visibility that information integration gives, but it is also a function of the discipline that ERP demands. Yet this discipline is itself a 'double-edged' sword. On one hand, it 'sharpens up' the management of every process within an organisation, allowing best practice (or at least common practice) to be implemented uniformly through the business. No longer will individual idiosyncratic behaviour by one part of a company's operations cause disruption to all other processes. On the other hand, it is the rigidity of this discipline that is both difficult to achieve and (arguably) inappropriate for all parts of the business. Nevertheless, the generally accepted benefits of ERP are as follows:

- Greater visibility of what is happening in all parts of the business.
- Forcing the business process-based changes that potentially make all parts of the business more efficient.
- Improved control of operations that encourages continuous improvement (albeit within the confines of the common process structures).
- More sophisticated communication with customers, suppliers and other business partners, often giving more accurate and timely information.
- Integrating whole supply chains including suppliers' suppliers and customers' customers.

Web-based and cloud-based ERP

An important justification for embarking on ERP is the potential it gives to link up with the outside world – for example, by integrating external communication systems into its internal ERP systems. However, as has been pointed out by some critics of the ERP software companies, different types of external company often need different types of information. Customers need to check the progress of their orders and invoicing, whereas suppliers and other partners want access to the details of operations planning and control. Not only that, but they want access all the time.

Web-based ERP systems provide access to the various modules of a commercial ERP system over the internet. Data and software can be located on an ERP vendor's servers as opposed to keeping everything on an operation's internal servers. The system is accessed via a web browser. While similar to web-based ERP, cloud-based ERP systems make use of cloud computing platforms and services. Cloud-based systems are usually located on the ERP vendor's servers with the user operation paying for the service on a 'per usage' basis, or for a time-based fee. This allows such systems to be scaled up (or down) as planning and control activities change. However, because they are usually based on a standard ERP template, cloud-based systems are not always capable of being highly customised.

Supply network ERP

The step beyond integrating internal ERP systems with immediate customers and suppliers is to integrate it with the systems of other businesses throughout the supply network. This is no easy task, in fact it is often exceptionally complicated. Not only do different ERP systems have to communicate with each other, they have to integrate with other types of systems. For example, sales and marketing functions often use systems such as customer relationship management (CRM) systems that manage the complexities of customer requirements, promises and transactions. Getting ERP and CRM systems to work together is itself often difficult. Nevertheless, such web-integrated ERP, or 'c-commerce' (collaborative commerce) applications are emerging and starting to make an impact on the way companies do business. Although a formidable task, the benefits are potentially great. The costs of communicating between supply network partners could be dramatically reduced and the potential for avoiding errors as information and products move between partners in the supply chain is significant. Yet such transparency also brings risks. If the ERP system of one operation within a supply chain fails for some reason, it may block the effective operation of the whole integrated information system throughout the network.

Case example

It's not that easy[3]

ERP systems have been called the nerve centre of an operation, and like actual nervous systems, they are difficult to deal with and can cause severe pain when they go wrong. And ERP implementation can go wrong, even when undertaken by experienced professionals. Look at these examples.

Lidl

Like many large companies, the German supermarket chain Lidl had put up with its in-house inventory management system that, after being in place for many years, was starting to creak. It commissioned SAP, the expert enterprise software firm, to install a totally new system. However, during the implementation process, it became clear that there was a clash between how Lidl preferred to account for inventory (at purchase price) and how most retailers do it (at the retail price they sell the goods for). SAP's system was set up for the latter. Because Lidl didn't want to change its accounting practice, SAP had to try to customise its system. This resulted in a series of implementation problems. What would already have been a complex project was not helped by staff turnover in Lidl's IT department. In 2018, seven years after starting, and after spending nearly €500 million, the project was scrapped.

Woolworth's Australia

Another retailer, this time in Australia, also had significant ERP implementation problems. Woolworths is Australia's largest supermarket chain (not to be confused with the defunct UK operation). Operating almost 1,000 stores across Australia, it employs 115,000 staff in its stores and supply networks. Its new ERP system was intended to modernise the company's planning and control efforts, but when, after six years of planning, the new system went live, problems emerged almost immediately. The most obvious symptom that something was wrong was the empty shelves in many of the company's stores. Apparently, a malfunction in the new system prevented the company from placing orders with its (many) suppliers. The investigation into what went wrong showed that one of the main problems was that during the ERP implementation process, insufficient attention was given to understanding and documenting the processes that were actually used by staff in their day-to-day running of the business. Much of the detail of these processes was just in the staff's heads, rather than formally recorded. So, when individuals left the company, they took key pieces of information with them. In effect, the ERP implementation failed because of a loss of corporate memory.

Oriola Finland

Oriola Finland is a Finnish health and well-being company with a strong position in the Swedish and Finnish healthcare markets. In Sweden, Oriola owns Kronans Apotek, the third-largest pharmacy chain in the country. It is a major distributor of health and well-being products that employs just under 3,000 people. But its ability to deliver its pharmaceuticals was severely compromised when a major ERP system upgrade went wrong. Oriola Finland deliver thousands of medications to pharmacists around the country, including insulin, cancer medications and anti-psychotics, so any disturbance in the supply chain doesn't just cause lost sales, it can damage people's health. This was a particularly serious disruption, given that 46 per cent of all drugs sold in Finland were supplied by Oriola, and switching to an alternative distributor would be difficult to do quickly. The problem was partly one of failure to sufficiently anticipate the disruption that any IT system change can cause. It is believed that Oriola did not anticipate any major supply disruption as the switchover happened, yet its ordering system was inoperative for days. The lesson is that, when transitioning to a new ERP, it is usually best to always plan for the worst. Eventually, Oriola did manage to sort out the problems with its new ERP, and also hired extra staff to process the backlog of orders. But the incident cost the company millions of euros and caused damage to its reputation for supply reliability.

Waste Management, Inc.

ERP implementation failures can sometimes end up in the law courts. Waste Management, Inc. is the leading provider of waste and environmental services in North America. When it announced that it was suing its ERP supplier, SAP, over the failure of an ERP implementation, it said that it was seeking the recovery of more than $100 million in project expenses, as well as the savings and benefits that the SAP software was promised to deliver. It said that SAP promised that the software could be fully implemented throughout all of Waste Management within 18 months, and that its software was an 'out-of-the-box' solution that would meet Waste Management's needs without any customisation or enhancements. However, according to Waste Management, the SAP implementation team discovered significant 'gaps' between the software's functionality and Waste Management's business requirements. Waste Management has discovered that these gaps were already known to the product development team in Germany, even before the SLA (service-level agreement) was signed. But members of SAP's implementation team had reportedly blamed Waste Management for the functional gaps and had submitted change orders requiring that Waste Management pay for fixing them. Five years after the complaint, the dispute was settled when SAP made a one-time cash payment to Waste Management.

Sales and operations planning (S&OP)

One of the problems with traditional operations planning and control is that, although several functions are often routinely involved in the process, each function could have a very different set of objectives. For example, marketing could be interested in maximising revenues and ensuring continuity of delivery to customers. Operations are likely to be under pressure to minimise costs (perhaps achieved through relatively long and stable operating levels). Finance will be interested in reducing working capital and inventory, and also reducing fixed costs. And so on. Yet these, and other functions such as engineering or human resources management, are all impacted by operations planning decisions and are probably involved in their own planning processes that partly depend on the output from the operations planning process. Sales and operations planning (S&OP) was first promoted as an important element of planning when materials requirements planning (see the supplement to this chapter) became a commonly used process. Early manufacturing resource planning implementations were often made less effective by the system being driven by unachievable plans. This is the dilemma that sales and operations planning (S&OP) is intended to address. It is a planning process that attempts to ensure that all tactical plans are aligned across the business's various functions and with the company's longer-term strategic plans.

It is a formal business process that looks over a period of 18–24 months ahead. In other words, it is not a purely short-term process. In fact, S&OP developed as an attempt to integrate short- and longer-term planning, as well as integrating the planning activities of key functions. It is an aggregated process that does not deal with detailed activities, but rather focuses on the overall (often aggregated) volume of output. Generally, it is a process that happens monthly, and tends to take place at a higher level, involving more senior management than traditional operations planning. S&OP also goes by many other names. It can be called integrated business planning, integrated business management, integrated performance management, rolling business planning, and regional business management, to name a few. It has also been noted[4] that some organisations continue to use the phrase 'S&OP', although they may mean something quite different.

10.3 Diagnostic question: *Are core planning and control activities effective?*

All resource planning and control activity eventually relies on a set of calculations that guide how much work to load onto different parts of the operation, when different activities should be performed, in what order individual jobs should be done, and how processes can be adjusted if they have deviated from plan. These calculations can be thought of as the 'engine room' of the whole resource planning and control system. Although the algorithms that guide the calculations are often embedded within computer-based systems, it is worthwhile understanding some of the core ideas on which they are based. These fall into four overlapping categories: loading, sequencing, scheduling, and monitoring and control.

Loading

Loading is the amount of work that is allocated to a process or stage or whole process. It is a capacity-related issue that will attempt to reconcile how much the operation or the process is expected to do with how much the operation or process can actually do. Essentially the loading activity calculates the consequences on individual parts of the operation's overall workload. It may or may not assume realistic capacity limits on what can be loaded. If it does, it is called finite loading; if not, it is called infinite loading. Finite loading is an approach that only allocates work to a work centre (a person, a machine, or perhaps a group of people or machines) up to a set limit. This limit is the estimate of capacity for that work centre (based on the times available for loading). Work over and above this capacity is not accepted. Figure 10.6(a) shows that the load on the work centre is not allowed to exceed the capacity limit. Finite loading is particularly relevant for operations where:

- *it is possible to limit the load* – for example, it is possible to run an appointment system for a general medical practice or a hairdresser;
- *it is necessary to limit the load* – for example, for safety reasons only a finite number of people and weight of luggage are allowed on aircraft;

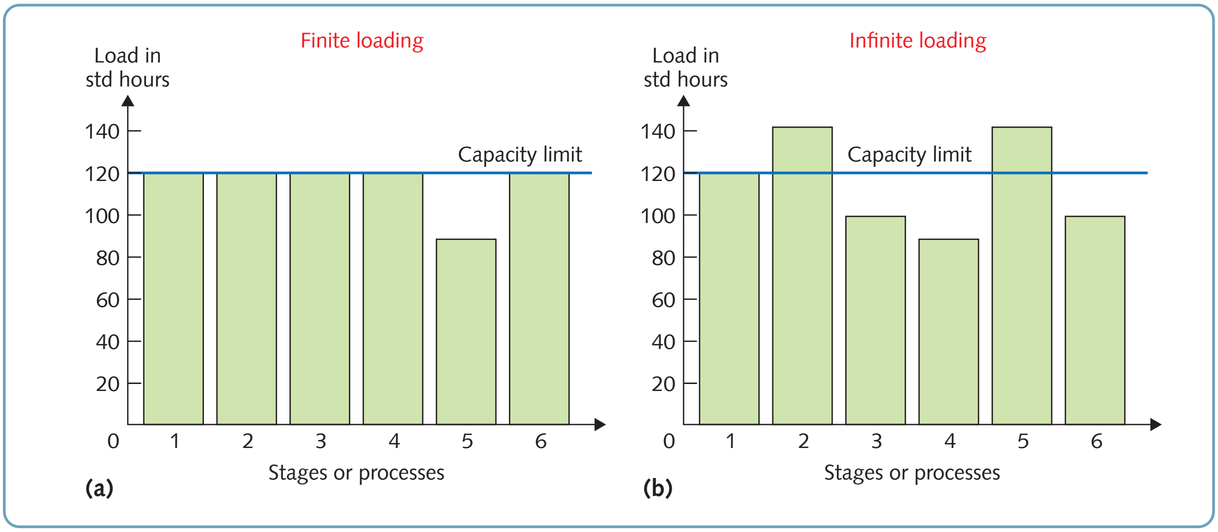


Figure 10.6 **Finite and infinite loading**

> OPERATIONS PRINCIPLE
> ***For any given level of demand, a planning and control system should be able to indicate the implications for the loading on any part of the operation.***

- *the cost of limiting the load is not prohibitive* – for example, the cost of maintaining a finite order book at a specialist sports car manufacturer does not adversely affect demand, and may even enhance it.

Infinite loading is an approach to loading work that does not limit accepting work, but instead tries to cope with it. Figure 10.6(b) illustrates a loading pattern where capacity constraints have not been used to limit loading. Infinite loading is relevant for operations where:

- *it is not possible to limit the load* – for example, an accident and emergency department in a hospital should not turn away arrivals needing attention;
- *it is not necessary to limit the load* – for example, fast-food outlets are designed to flex capacity up and down to cope with varying arrival rates of customers and during busy periods, customers accept that they must queue for some time before being served;
- *the cost of limiting the load is prohibitive* – for example, if a retail bank turned away customers at the door because a set amount were inside, customers would feel less than happy with the service.

In complex planning and control activities where there are multiple stages each with different capacities, and with a varying mix arriving at the facilities, such as a machine shop in an engineering company, the constraints imposed by finite loading make loading calculations complex and not worth the considerable computational power that would be needed.

Sequencing

After the 'loading' of work onto processes, the order or sequence in which it will be worked on needs to be determined. This task is called 'sequencing'. The priorities given to work in an operation are often determined by some predefined set of sequencing rules. Some of these are summarised here.

Customer priority – This allows an important or aggrieved customer, or item, to be prioritised, irrespective of their order of arrival. Some banks, for example, give priority to important customers. Accident and emergency departments in hospitals must rapidly devise a schedule that prioritises patients presenting symptoms of a serious illness. Hospitals have developed 'triage systems', whereby medical staff assess patients briefly on arrival to determine their relative urgency.

Due date (DD) – Work is sequenced according to when it is 'due' for delivery, irrespective of the size of each job or the importance of each customer. For example, a support service in an office block, such as a reprographic unit, may sequence the work according to when the job is needed. Due-date sequencing usually improves delivery reliability and average delivery speed, but may not provide optimal productivity.

Last in, first out (LIFO) – This is usually selected for practical reasons. For example, unloading an elevator is more convenient on a LIFO basis, as there is only one entrance and exit. LIFO has a very adverse effect on delivery speed and reliability.

First in, first out (FIFO) – Also called 'first come, first served' (FCFS), this is a simple and equitable rule, used especially when queues are evident to customers – for example in theme parks.

Longest operation time first (LOT) – Executing the longest job first has the advantage of utilising work centres for long periods but, although utilisation may be high (therefore cost relatively low), this rule does not take into account delivery speed, delivery reliability or flexibility.

Shortest operation time first (SOT) – This sends small jobs quickly through the process, achieving output quickly and enabling revenue to be generated quickly. Short-term delivery performance may be improved, but productivity and the throughput time of large jobs is likely to be poor.

Case example

The trials of triage[5]

Dominique-Jean Larrey was a military surgeon in the French army during Napoleon's campaigns. When treating battlefield casualties, Larrey had to decide (regardless of their military rank) which soldiers needed medical attention most urgently. To help in what was a difficult decision, both medically and ethically, he developed the concept of distinguishing between urgent and non-urgent patients. This became known as 'triage', from the French *trier* (to separate out). It is still used today to sequence patients waiting for limited medical resources such as staff, beds, intensive care and ventilators. The main difference is that, in Napoleonic times, many injuries meant certain death. Now, technological progress means that most can be treated – if the resources are available. For example, hospital accident and emergency departments have limited staff, beds and equipment, and patients arrive at random. It is then the job of the hospital's reception and medical staff to assess and prioritise patients according to the relative seriousness of their illness or injury. Patients with very serious injuries, or presenting symptoms of a serious illness, generally need to be attended to urgently. Patients in some discomfort, but whose injuries or illnesses are not life-threatening, will have to wait until the urgent cases are treated. Routine non-urgent cases will have the lowest priority. In many circumstances, these patients will have to wait many hours if the hospital is busy. They may even be turned away if the hospital is too busy with more important cases.

Triage always presents moral dilemmas and trade-offs. This was especially exposed during the COVID-19 pandemic. Doctors were faced with the most appalling decisions about how to allocate scarce resources. Under extreme resource shortage, it may not be the patients who are the most sick who are given priority access to resources. In fact, beyond a certain point, a severely sick patient may, under some circumstances, be less likely to receive treatment. It can be argued that, especially if treatment is likely to be distressing to the patient, and a person is inevitably going to die, it is simply cruel to add to their affliction by, for example, putting them on a ventilator. Moreover, the equipment could be better used for a patient who is less ill but more likely to survive. Making such trade-offs takes its toll on medical staff. During the pandemic there were reports of doctors weeping in hospital corridors because of the choices they had to make. Although most professionals agreed that resources were allocated to the patients who had the greatest chances of successful treatment, and who had the greatest life expectancy, it still meant having to take some brutal decisions that literally meant life or death. Doctors have reported that it helps if the criteria and decision framework for distinguishing between patients is decided in advance, and patients and their families are carefully and sensitively informed. It may also be better for someone other than front-line doctors to make the difficult decisions. Some states in the USA have triage officers or committees that make such decisions, with front-line doctors free to appeal a decision if they think it is mistaken.


Shutterstock

Worked example

Steve Smith, website designer

Returning from his annual vacation (he finished all outstanding jobs before he left), five design jobs are given to Steve upon his arrival at work. He gives them the codes A to E. Steve has to decide in which sequence to undertake the jobs. He wants both to minimise the average time the jobs are tied up in his office and, if possible, to meet the deadlines (delivery times) allocated to each job.

His first thought is to do the jobs in the order they were given to him, i.e. first in, first out (FIFO):

Sequencing rule – first in, first out (FIFO)

Sequence of jobs	*Process time (days)*	*Start time*	*Finish time*	*Due date*	*Lateness (days)*
A	5	0	5	6	0
B	3	5	8	5	3
C	6	8	14	8	6
D	2	14	16	7	9
E	1	16	17	3	14
Total time in process		60	**Total lateness**		32
Average time in process (total/5)		**12**	**Average lateness (total/5)**		**6.4**

Alarmed by the average lateness, he tries the due date (DD) rule:

Sequencing rule – due date (DD)

Sequence of jobs	*Process time (days)*	*Start time*	*Finish time*	*Due date*	*Lateness (days)*
E	1	0	1	3	0
B	3	1	4	5	0
A	5	4	9	6	3
D	2	9	11	7	4
C	6	11	17	8	9
Total time in process		**42**	**Total lateness**		**16**
Average time in process (total/5)		**8.4**	**Average lateness (total/5)**		**3.2**

Better! But Steve tries out the shortest operation time first (SOT) rule:

Sequencing rule – shortest operation time first (SOT)

Sequence of jobs	*Process time (days)*	*Start time*	*Finish time*	*Due date*	*Lateness (days)*
E	1	0	1	3	0
D	2	1	3	7	0
B	3	3	6	5	1
A	5	6	11	6	5
C	6	11	17	8	9
Total time in process		**38**	**Total lateness**		**16**
Average time in process (total/5)		**7.6**	**Average lateness (total/5)**		**3.2**

This gives the same degree of average lateness but with a lower average time in the process. Steve decides to use the SOT rule.

There are many sequencing rules of differing complexity. Although different rules will perform differently depending on the circumstances of the sequencing problem, in practice the SOT rule generally performs well.

Case example

Can airline passengers be sequenced?[6]

Like many before him, Dr Jason Steffen, a professional astrophysicist from the world famous Fermilab, was frustrated by the time it took to load him and his fellow passengers onto an aircraft. He decided to devise a way to make the experience a little less tedious. So, for a while, he neglected his usual work of examining extra-solar planets, dark matter and cosmology, and experimentally tested a faster method of boarding aircraft. He found that, by changing the sequence in which passengers are loaded onto the aircraft, airlines could potentially save both time and money. Using a computer simulation and the arithmetic techniques routinely used in his day-to-day work, he was able to find what seemed to be a superior sequencing method. In fact, the most common way of boarding passenger planes proved to be the least efficient. This is called the 'block method' where blocks of seats are called for boarding, starting from the back. Previously, other experts in the airline industry had suggested boarding those in window seats first followed by middle and aisle seats. This is called the Wilma method. But according to Dr Steffen's simulations, two things slow down the boarding process. The first is that passengers may be required to wait in the aisle while those ahead of them store their luggage before they can take their seat. The second is that passengers already seated in aisle or middle seats frequently have to rise and move into the aisle to let others take seats nearer the window. So Dr Steffen suggested a variant of the Wilma method that minimised the first type of disturbance and eliminated the second. He suggested boarding in alternate rows, progressing from the rear forward, window seats first. Using this approach (now called the Steffen method), first the window seats for every other row on one side of the plane are boarded. Next, alternate rows of window seats on the opposite side are boarded. Then, the window seats in the skipped rows are filled in on each side. The procedure then repeats with the middle seats and the aisles (see Figure 10.7).

Later, the effectiveness of the various approaches was tested using a mock-up of a Boeing 757 aircraft and 72 luggage-carrying volunteers. Five different scenarios were tested: block boarding in groups of rows from back to front; one by one from back to front; the Wilma method; the Steffen method; and completely random boarding. In all cases, parent–child pairs were allowed to board first as it was assumed that families were likely to want to stay together. As Dr Steffen had predicted, the conventional block approach came out as the slowest, with the strict back-to-front approach not much better. Completely random boarding (unallocated seating), which is used by several low-cost airlines, fared much better, most probably

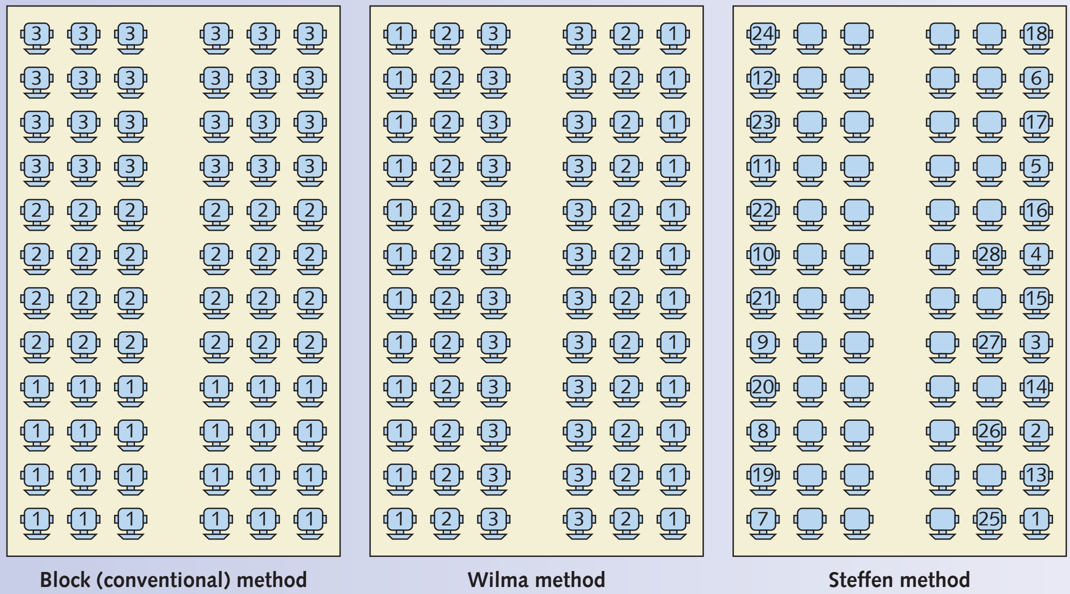


Figure 10.7 Calculating the best way to sequence passengers onto an aircraft

because it randomly avoids space conflicts. The times for fully boarding the 72 passengers using each method were as follows: block boarding – 6:54 minutes; back-to-front – 6:11 minutes; random boarding – 4:44 minutes; Wilma method – 4:13 minutes; Steffen method – 3:36 minutes.

The big question is, would passengers really be prepared to be sequenced in this way as they queue to board an aircraft? Some airlines argue that directing passengers onto a plane is a little like herding cats. But if they could be persuaded to adopt Dr Steffen's system it would save time for customers and very significant amounts of money for airlines.

Imagedoc/Alamy Stock Photo

Scheduling

Scheduling is the activity of producing a detailed timetable showing when activities should start and end. Schedules are familiar in many consumer environments – for example, a bus schedule that shows the time each bus is due to arrive at each stage of the route. But, although familiar, it is one of the most complex tasks in operations and process management. Schedules may have to deal simultaneously with many activities and several different types of resources, probably with different capabilities and capacities. Also, the number of possible schedules increases rapidly as the number of activities and resources increase. If one process has five different jobs to process, any of the five jobs could be processed first and, following that, any one of the remaining four jobs, and so on. This means that there are: $5 \times 4 \times 3 \times 2 = 120$ different schedules possible. More generally, for n jobs there are $n!$ (factorial n, or $n \times (n-1) \times (n-2) \ldots \times 1$) different ways of scheduling the jobs through a single process or stage. If there is more than

OPERATIONS PRINCIPLE
An operation's planning and control system should allow for the effects of alternative schedules to be assessed.

one process or stage, there are $(n!)^m$ possible schedules, where n is the number of jobs and m is the number of processes or stages. In practical terms, this means that there are often many millions of feasible schedules, even for relatively small operations. This is why scheduling rarely attempts to provide an 'optimal' solution but rather satisfies itself with an 'acceptable' feasible one.

Gantt charts

The most common method of scheduling is by use of the Gantt chart. This is a simple device that represents time as a bar, or channel, on a chart. The start and finish times for activities can be indicated on the chart and sometimes the actual progress of the job is also indicated. The advantages of Gantt charts are that they provide a simple visual representation both of what should be happening and of what actually is happening in the operation. Furthermore, they can be used to 'test out' alternative schedules. It is a relatively simple task to represent alternative schedules (even if it is a far from simple task to find a schedule that fits all the resources satisfactorily). Figure 10.8 illustrates a Gantt chart for a specialist software developer. It indicates the progress of several jobs as they are expected to progress through five stages of the process. Gantt charts are not an optimising tool, they merely facilitate the development of alternative schedules by communicating them effectively.

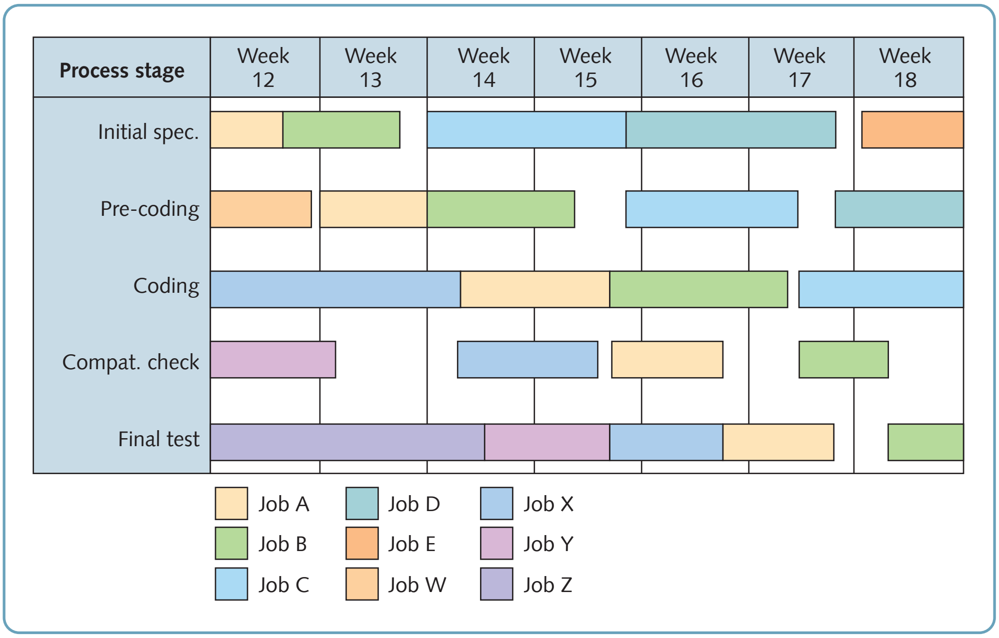


Figure 10.8 **Gantt chart showing the schedule for jobs at each process stage**

Case example

Sequencing and scheduling at London's Heathrow Airport[7]

Heathrow is the UK's busiest airport, and the busiest two-runway airport in the world, welcoming around 1,300 combined take-offs and landings each day. Landing around 650 aircraft in a day, air traffic controllers have one of the most complex sequencing jobs to perform, as they decide which aircraft to call down next from their waiting areas (known as 'stacks') to land on one of the two runways. Many airports use a sequencing policy based on 'first come, first served'. However, this does not always give the best airport performance, where performance is assessed by such measures as runway utilisation, total aircraft throughput, passenger throughput and passenger waiting time. For very busy airports like Heathrow, a more sophisticated sequencing approach is needed. For most of the time at Heathrow, one runway is used solely for take-offs and the other solely for landings (known as a 'segregated' operating mode). However, at particularly busy times, both runways can be used for landings. Safety considerations are, of course, paramount in deciding on an appropriate sequence. There must be a minimum time and distance between aircraft when they take off or land. This is because of what is known as the 'wake vortex' – turbulence that is caused by the 'lift' component of flight (without which the aircraft could not fly). Lift is caused by the pressure difference between the upper and lower surfaces of the wing. Wake vortices can result in turbulent conditions if an aircraft follows too close to the previous one, which passengers would find uncomfortable, and possibly distressing. It could even cause possible damage to the following aircraft. The magnitude of a wake vortex depends on the size of the aircraft, large aircraft causing more air turbulence. So, following a large aircraft means leaving a (relatively) long time delay before another aircraft can land. Conversely, a light aircraft generates little air turbulence and therefore only a (relatively) short time delay is needed before other aircraft can land. In other words, the sequence in which aircraft are called to land will determine the total time taken to complete landing. But, in addition to deciding the sequence in which aircraft will land, controllers must also construct a schedule that determines a landing time for each aircraft. This schedule should:

- Allow sufficient time for an aircraft to fly safely from its current position in the stack to the runway, so that it will land at the appropriate position in the sequence.
- Make sure that any aircraft do not run low on fuel while waiting to land.
- Ensure that aircraft do not land too close together.

To complicate the task further, weather can also influence decisions. Aircraft have to take off and land against the wind, so the landing direction depends on the prevailing wind (which can change). Which is why meteorological experts are constantly monitoring weather conditions prevailing at 30,000 feet. The airport must also try to minimise the noise nuisance caused to local communities, which means that no landings are allowed before 04.30, with a maximum of 16 flights before 06.00, and preferably the quietest planes.

Scheduling work patterns

Where the dominant resource in an operation is its staff, then the schedule of work times effectively determines the capacity of the operation itself. Scheduling needs to make sure enough people are working at any time to provide a capacity appropriate for the level of demand. Operations such as call centres and hospitals, which must respond directly to customer demand, will need to schedule the working hours of their staff with demand in mind. For example, Figure 10.9 shows the scheduling of shifts for a small technical 'hotline' support service in a software company. Its service times are 04.00 to 20.00 on Monday, 04.00 to 22.00 Tuesday to Friday, 06.00 to 22.00 on Saturday, and 10.00 to 20.00 on Sunday. Demand is heaviest Tuesday to Thursday, starts to decrease on Friday, is low over the weekend and starts to increase again on Monday. The scheduling task for this kind of problem can be considered over different timescales, two of which are shown in Figure 10.9. During the day, working hours need to be agreed with individual staff members. During the week, days off need to be agreed. During the year, vacations, training periods and other blocks of time where staff are unavailable need to be agreed. All this has to be scheduled such that:

- capacity matches demand;
- the length of each shift is neither excessively long nor too short to be attractive to staff;
- working at unsocialable hours is minimised;
- days off match agreed staff conditions – in this example, staff prefer two consecutive days off every week;
- vacation and other 'time-off' blocks are accommodated;
- sufficient flexibility is built into the schedule to cover for unexpected changes in supply (staff illness) and demand (surge in customer calls).

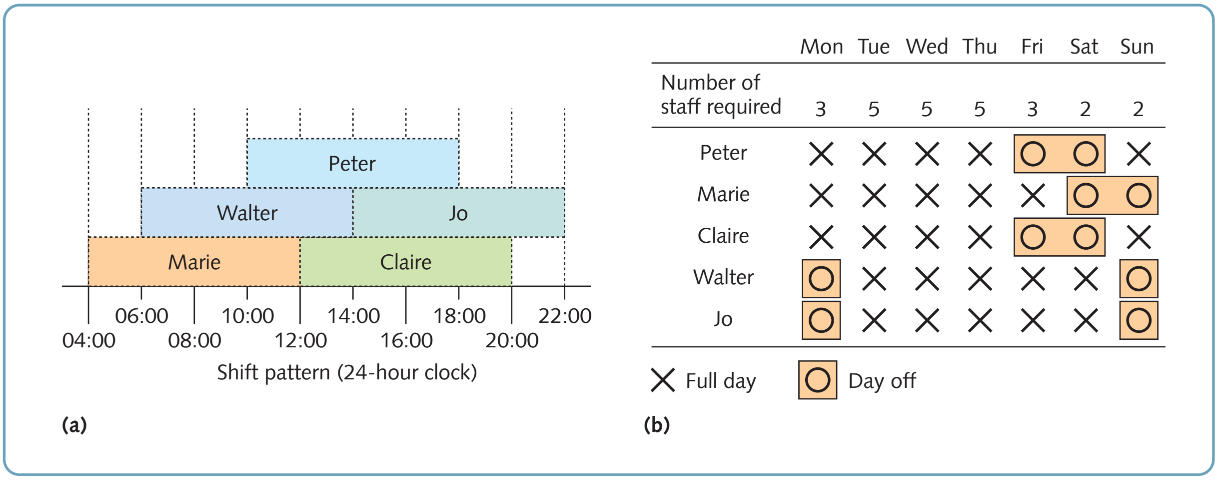


Figure 10.9 Shift allocation for the technical 'hotline' (a) on a daily basis, (b) on a weekly basis

Scheduling staff times is one of the most complex of scheduling problems. In the relatively simple example shown in Figure 10.9 we have assumed that all staff have the same level and type of skill. In very large operations with many types of skills to schedule and uncertain demand (for example a large hospital), the scheduling problem becomes extremely complex. Some mathematical techniques are available but most scheduling of this type is, in practice, solved using heuristics (rules of thumb), some of which are incorporated into commercially available software packages.

Theory of constraints (TOC)

An important concept, closely related to scheduling, which recognises the importance of planning to known capacity constraints, is the theory of constraints (TOC). It focuses scheduling effort on the bottleneck parts of the operation. By identifying the location of constraints, working to remove them, and then looking for the next constraint, an operation is always focusing on the part that critically determines the pace of output. The approach that uses this idea is called optimised production technology (OPT). Its development and the marketing of it as a proprietary software product were originated by Eliyahu Goldratt.[8] It helps to schedule production systems to the pace dictated by the most heavily loaded resources – that is, bottlenecks. If the rate of activity in any part of the system exceeds that of the bottleneck, then items are being produced that cannot be used. If the rate of working falls below the pace at the bottleneck, then the entire system is underutilised. The 'principles' of underlying OPT demonstrate this focus on bottlenecks.

OPT principles

1. Balance flow, not capacity. It is more important to reduce throughput time rather than achieving a notional capacity balance between stages or processes.

2. The level of utilisation of a non-bottleneck is determined by some other constraint in the system, not by its own capacity. This applies to stages in a process, processes in an operation and operations in a supply network.

3. Utilisation and activation of a resource are not the same. According to the TOC a resource is being *utilised* only if it contributes to the entire process or operation creating more output. A process or stage can be *activated* in the sense that it is working, but it may only be creating stock or performing some other non-value-added activity.

4. An hour lost (not used) at a bottleneck is an hour lost forever out of the entire system. The bottleneck limits the output from the entire process or operation, therefore the underutilisation of a bottleneck affects the entire process or operation.
5. An hour saved at a non-bottleneck is a mirage. Non-bottlenecks have spare capacity anyway. Why bother utilising them even less?
6. Bottlenecks govern both throughput and inventory in the system. If bottlenecks govern flow, then they govern throughput time, which in turn governs inventory.
7. You do not have to transfer batches in the same quantities as you produce them. Flow will probably be improved by dividing large production batches into smaller ones for moving through a process.
8. The size of the process batch should be variable, not fixed. Again, from the EBQ model (see Chapter 9), the circumstances that control batch size may vary between different products.
9. Fluctuations in connected and sequence-dependent processes add to each other rather than averaging out. So, if two parallel processes or stages are capable of a particular average output rate, in parallel, they will never be able to produce the same average output rate.
10. Schedules should be established by looking at all constraints simultaneously. Because of bottlenecks and constraints within complex systems, it is difficult to work out schedules according to a simple system of rules. Rather, all constraints need to be considered together.

Monitoring and control

Having created a plan for the operation through loading, sequencing and scheduling, each part of the operation has to be monitored to ensure that activities are happening as planned. Any deviation can be rectified through some kind of intervention in the operation, which itself will probably involve some replanning. Figure 10.10 illustrates a simple view of control. The output from a work centre is monitored and compared with the plan that indicates what the work centre is supposed to be doing. Deviations from this plan are taken into account through a replanning activity and the necessary interventions made (on a timely basis) to the work centre, which will ensure that the new plan is carried out. Eventually, however, some further deviation from planned activity will be detected and the cycle is repeated.

> **OPERATIONS PRINCIPLE**
> *A planning and control system should be able to detect deviations from plans within a timescale that allows an appropriate response.*

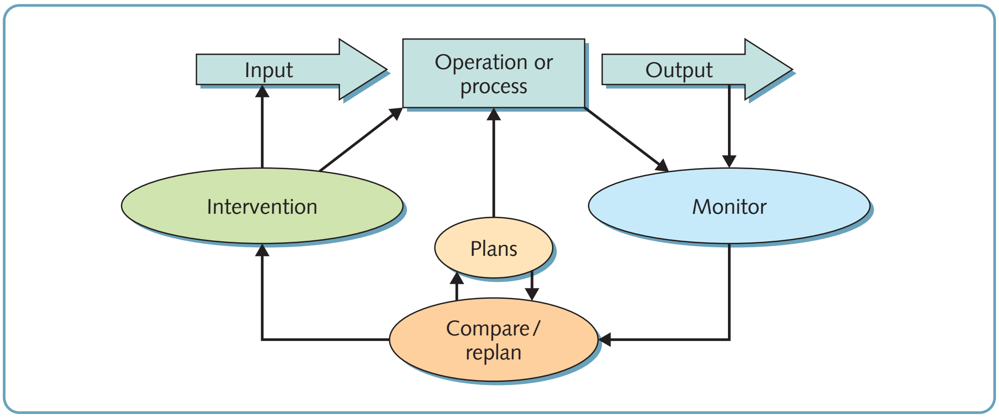


Figure 10.10 A simple model of control

Push and pull control

One element of control is periodic intervention into processes and operations. A key distinction is between intervention signals that push work through processes and operations and those that pull work only when it is required. In a *push* system of control, activities are scheduled by means of a central system and completed in line with central instructions, such as an MRP system (see the supplement to this chapter). Each work centre pushes out work without considering whether the succeeding work centre can make use of it. Deviations from plan are noted by the central operations planning and control system, and plans adjusted as required. In a *pull* system of control, the pace and specification of what is done are set by the succeeding 'customer' workstation, which 'pulls' work from the preceding (supplier) workstation. The customer acts as the only 'trigger' for movement. If a request is not passed back from the customer to the supplier, the supplier cannot produce anything or move any materials. A request from a customer not only triggers production at the supplying stage, but also prompts the supplying stage to request a further delivery from its own suppliers. In this way, demand is transmitted back through the stages from the original point of demand by the original customer.

Push systems of control are more formal and require significant decision-making or computing power when it is necessary to replan in the light of events. But push control can cope with very significant changes in circumstances such as major shifts in output level or product mix. By contrast, pull control is more self-adjusting in the sense that the rules that govern relationships between stages or processes cope with deviations from plan without reference to any higher decision-making authority. But there are limits to the extent that this can cope with major fluctuations in demand. Pull control works best when conditions are relatively stable. Understanding the differences between push and pull is also important because they have different effects in terms of their propensities to accumulate inventory. Pull systems are far less likely to result in inventory build-up and therefore have advantages in terms of the lean synchronisation of flow (covered in Chapter 11).

OPERATIONS PRINCIPLE
Pull control reduces the build-up of inventory between processes or stages.

Drum, buffer, rope control

The drum, buffer, rope concept comes from the theory of constraints (TOC) described earlier. It is an idea that helps to decide exactly *where* control should occur. Again, the TOC emphasises the role of the bottleneck on workflow. If the bottleneck is the chief constraint, it should be the control point of the whole process. The bottleneck should be the *drum* because it sets the 'beat' for the rest of the process to follow. Because it does not have sufficient capacity, a bottleneck is (or should be) working all the time. Therefore, it is sensible to keep a *buffer* of inventory in front of it to make sure that it always has something to work on. Also, because it constrains the output of the whole process, any time lost at the bottleneck will affect the output from the whole process. So it is not worthwhile for the parts of the process before the bottleneck to work to their full capacity. All they would do is produce work that would accumulate further along in the process up to the point where the bottleneck is constraining flow. Therefore, some form of communication between the bottleneck and the input to the process is needed to make sure that activities before the bottleneck do not overproduce. This is called the *rope* (see Figure 10.11).

OPERATIONS PRINCIPLE
The constraints of bottleneck processes and activities should be a major input to the planning and control activity.

Controlling operations is not always routine

The simple monitoring-control model in Figure 10.10 helps us to understand the basic functions of the monitoring and control activity. But it is a simplification. Some simple routine processes may approximate to it, but many other operations do not. In fact, some of the specific criticisms

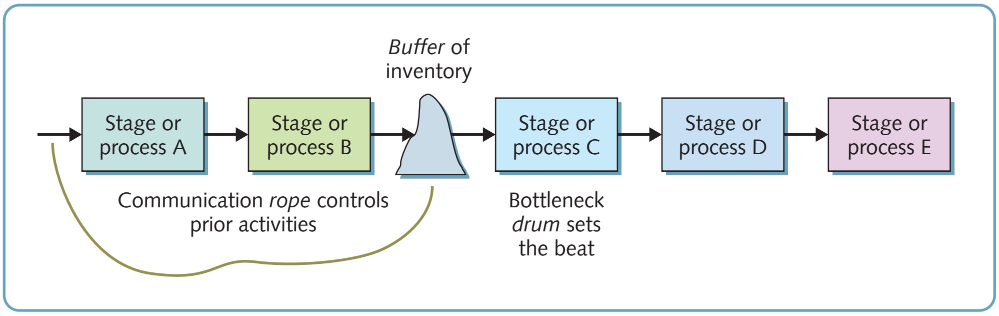


Figure 10.11 **The drum, buffer, rope concept**

provide a useful set of questions that can be used to assess the degree of difficulty associated with control of any operation. In particular:

- Is there consensus over what the operation's objectives should be?
- Are the effects of interventions into the operation predictable?
- Are the operation's activities largely repetitive?

Starting with the first question, are strategic objectives clear and unambiguous? It is not always possible (or necessarily desirable) to articulate every aspect of an operation's objectives in detail. Many operations are just too complex for that. Nor does every senior manager always agree on what the operation's objectives *should* be. Often the lack of a clear objective is because individual managers have different and conflicting interests. In social care organisations for example, some managers are charged with protecting vulnerable members of society, others with ensuring that public money is not wasted, and yet others may be required to protect the independence of professional staff. At other times, objectives are ambiguous because the strategy has to cope with unpredictable changes in the environment, making the original objectives redundant. A further assumption in the simplified control model is that there is some reasonable knowledge of how to bring about the desired outcome. That is, when a decision is made, one can predict its effects with a reasonable degree of confidence. In other words, operational control assumes that any interventions that are intended to bring a process back under control will indeed have the intended effect. Yet, this implies that the relationships between the intervention and the resulting consequence within the process are predictable, which in turn assumes that the degree of process knowledge is high. For example, if an organisation decides to relocate in order to be more convenient for its customers, it may or may not prove to be a correct decision. Customers may react in a manner that was not predicted. Even if customers seem initially to respond well to the new location, there may be a lag before negative reactions become evident. In fact, many operations decisions are taken on activities about which the cause–effect relationship is only partly understood. The final assumption about control is that control interventions are made in a repetitive way and occur frequently (for example checking on a process hourly or daily), meaning that the operation has the opportunity to learn how its interventions affect the process, which considerably facilitates control. However, some control situations are non-repetitive; for example, those involving unique services or products. So, because the intervention, or the deviation from plan that caused it, may not be repeated, there is little opportunity for learning.

Figure 10.12 illustrates how these questions can form a 'decision tree' type of model that indicates how the nature of operations control may be influenced.[9] Operational control is relatively straightforward: objectives are unambiguous, the effects of interventions are known, and activities are repetitive. This type of control can be carried out using predetermined conventions and rules. There are, however, still some challenges to successful routine control. It needs operational discipline to make sure that control procedures are implemented systematically. The main point though is that any divergence from the conditions necessary for routine control implies that a different type of control will be needed.

OPERATIONS PRINCIPLE
Planning and control is not always routine, especially when objectives are ambiguous, the effects of interventions into the operation are not predictable and activities are not repetitive.

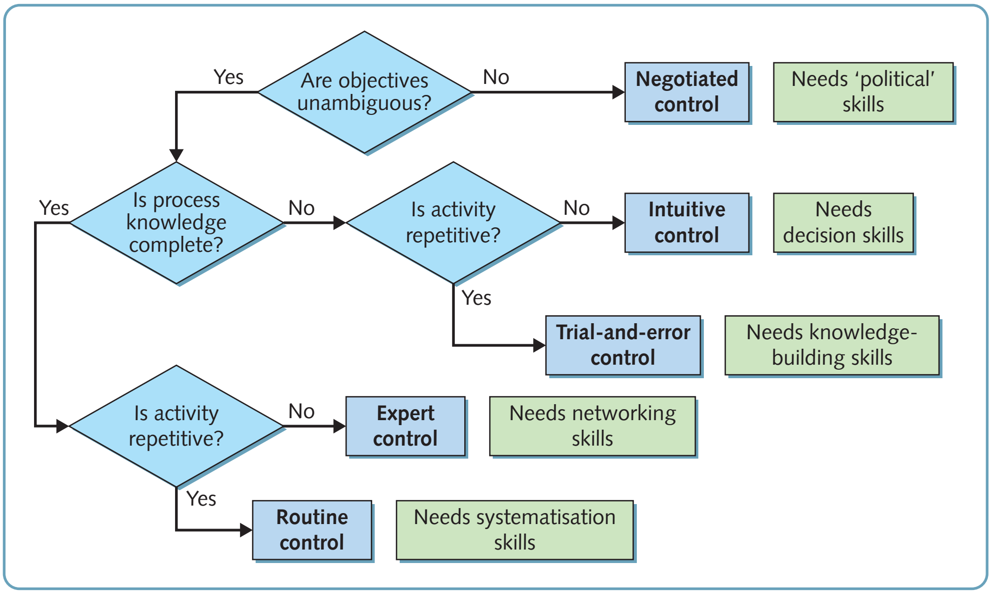


Figure 10.12 **Control is not always routine; different circumstances require different types of control**

Expert control – If objectives are unambiguous, and the effects of interventions relatively well understood, but the activity is not repetitive (for example, installing or upgrading software or IT systems), control can be delegated to an 'expert' – someone for whom such activities *are* repetitive because they have built their knowledge on previous experience elsewhere. Making a success of expert control requires that such experts exist and can be 'acquired' by the firm. It also requires that the expert takes advantage of the control knowledge already present in the firm and integrates his or her 'expert' knowledge with the support that potentially exists internally. Both of these place a stress on the need to 'network', both in terms of acquiring expertise and then integrating that expertise into the organisation.

Trial-and-error control – If strategic objectives are relatively unambiguous, but effects of interventions not known, yet the activity is repetitive, the operation can gain knowledge of how to control successfully through its own failures. In other words, although simple prescriptions may not be available in the early stages of making control interventions, the organisation can learn how to do it through experience. For example, if a fast-food chain is opening new stores into new markets, it may not be sure how best to arrange the openings at first. But if the launch is the first of several, the objective must be not only to make a success of each launch, but equally (or more) importantly, it must learn from each experience. It is these knowledge-building skills that will ultimately determine the effectiveness of trial-and-error control.

Intuitive control – If objectives are relatively unambiguous (so it is clear what the operation is trying to do), but effects of control interventions are not known, and nor are they repetitive, then learning by trial and error is not possible. Here control becomes more of an art than a science. And in these circumstances control must be based on the management team using its intuition to make control decisions. Many strategic operations processes fall into this category – for example setting up a strategic supply partnership (see Chapter 7). Objectives are clear (jointly survive in the long term, make an acceptable return, and so on), but not only are control interventions not repetitive and their effects not fully understood, sometimes the supplier's interests may be in conflict with yours. Yet, simply stating that 'intuition' is needed in these circumstances is not particularly helpful. Instincts and feelings are, of course, valuable attributes in any management team, but they are the result, at least partly, of understanding how best to organise any shared understanding, knowledge, and decision-making skills. Thorough decision analysis is required, not to make the decision 'mechanistically', but to frame it so that connections can be made, consequences understood and insights gained.

Negotiated control – The most difficult circumstance for strategic control is when objectives are ambiguous. This type of control involves reducing ambiguity in some way by making objectives less uncertain. Sometimes this is done simply by senior managers 'pronouncing' or arbitrarily deciding what objectives *should* be, irrespective of opposing views. For example, controlling the activities of a childcare service can involve very different views among the professional social workers making day-to-day decisions. Often a negotiated settlement may be sought, which then can become an unambiguous objective. Alternatively, outside experts could be used, either to help with the negotiations, or to remove control decisions from those with conflicting views. But, even within the framework of negotiation, there is almost always a political element when ambiguities in objectives exist. Negotiation processes will be, to some extent, dependent on power structures.

Critical commentary

- Far from being the magic ingredient that allows operations to fully integrate all their information, ERP is regarded by some as one of the most expensive ways of getting zero or even negative return on investment. For example, the American chemicals giant, Dow Chemical, spent almost half-a-billion dollars and seven years implementing an ERP system that became outdated almost as soon as it was implemented. One company, FoxMeyer Drug, claimed that the expense and problems that it encountered in implementing ERP eventually drove it into bankruptcy. One problem is that ERP implementation is expensive. This is partly because of the need to customise the system, understand its implications for the organisation, and train staff to use it. Spending on what some call the *ERP ecosystem* (consulting, hardware, networking and complementary applications) has been estimated as being twice the spending on the software itself. But it is not only the expense that has disillusioned many companies, it is also the returns they have had on their investment. Some studies show that the vast majority of companies implementing ERP are disappointed with the effect it has had on their businesses. Certainly, many companies find that they have to (sometimes fundamentally) change the way they organise their operations in order to fit in with ERP systems. This organisational impact of ERP (which has been described as the corporate equivalent of root-canal work) can have a significantly disruptive effect on the organisation's operations.
- If one accepts only some of the criticisms of ERP, it does pose the question as to why companies have invested such large amounts of money in it. Partly it is the attraction of turning the company's information systems into a 'smooth-running and integrated machine'. The prospect of such organisational efficiency is attractive to most managers, even if it does presuppose a very simplistic model of how organisations work in practice. After a while, although organisations could now see the formidable problems in ERP implementation, the investments were justified on the basis that, 'even if we gain no significant advantage by investing in ERP, we will be placed at a disadvantage by *not* investing in it because all our competitors are doing so'. There is probably some truth in this; sometimes businesses have to invest just to stand still.

SUMMARY CHECKLIST

- ☐ Is appropriate effort devoted to planning and controlling the operation's resources and activities?
- ☐ Have any recent failures in planning and control been used to reconsider how the planning and control system operates?
- ☐ Does the system interface with customers so as to encourage a positive customer experience?
- ☐ Does the planning and control system interface with suppliers so as to promote a supplier experience that is in your long-term interests?
- ☐ Does the system perform basic planning and control calculations in an appropriate and realistic manner?
- ☐ Is the balance between human and automated decision-making understood and appropriate for the circumstances?
- ☐ How well is resource planning and control information integrated?
- ☐ Have the advantages and disadvantages of moving to a sophisticated (but expensive!) ERP system been investigated?
- ☐ If so, have the possibilities of web integration and supply chain scope been investigated?
- ☐ Are bottlenecks accounted for in the way planning and control decisions are made?
- ☐ If not, have bottlenecks been identified and their effect on the smooth flow of items through the operation been evaluated?

Case study

Audall Auto Servicing

It had been ten years since Dan Audall founded Audall Auto Servicing as an independent vehicle-servicing and repair business. Previously he had been the manager of the servicing department of a 'premium' car dealership, the experience of which had convinced him that there was demand in the area for servicing one of the 'rival' makes of car. '*We were continually getting requests from owners to service or repair their vehicles, partly because we had a good reputation, but mainly because there were no local dealerships that could do that kind of work. Owners had to use small independents or travel a long distance to get to the nearest dealership. I persuaded the auto company that I could provide appropriate service for their vehicles without taking significant business away from their other franchised service centres. It was a gamble, but they backed me*.'

That was ten years ago, and Dan's gamble had paid off. Audall Auto Servicing had grown to the point where he had invested in a modern service centre close to his first location. The outline plans for the new centre are shown in Figure 10.13. It had five servicing bays, a parts store, a car wash and a customer waiting area. '*Although the new building is the same nominal capacity as the old one, it gives us more room for the technicians to move about, and the customer waiting area is a distinct improvement in the service that we can offer customers. There is also room next to this building that we could use for expansion, although I would prefer to wait a couple of years before committing to this investment. Before then, I think that we could do more business with our existing capacity*.' (Dan Audall)

Dan's conviction that the operation could do more business with its current capacity was based on performance figures covering the first two months in the new building, showing that the servicing bays were, on average, only 83 per cent utilised. This was a figure that Dan believed could be improved, as did Diya Chopra, his office manager. However, Diya thought that the room for improvement would be limited: '*No week is ever perfectly predictable. There are just too many uncertainties. Even*

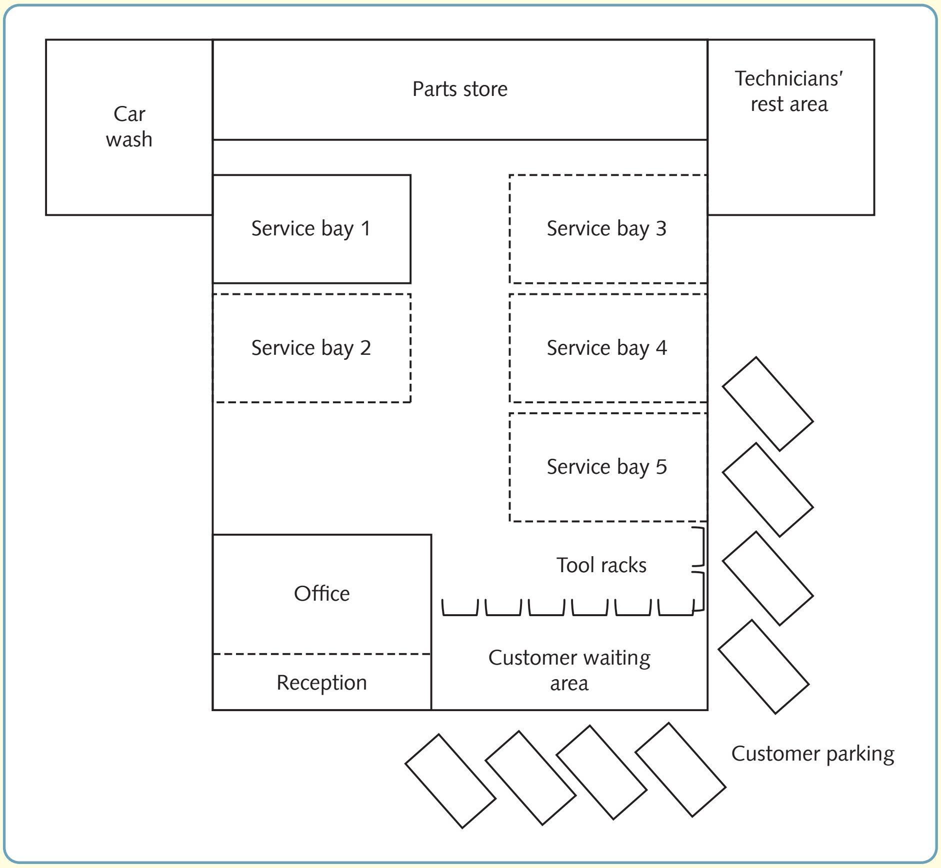


Figure 10.13 Outline plan of Audall Auto Servicing service centre

a minor service (usually every 6,000 miles or 10,000 km) can throw up problems that can take two or three times the time that we have allowed. Major services (usually every 12,000 miles or 18,000 km) are even less predictable. With these "standard" services, customers often want us to keep the vehicle until it's repaired rather than make another booking. But the most unpredictable are what we call the "short-term" repairs where a customer wants their vehicle "up and running" as soon as possible. We call these jobs short term not because they take little time, but because they are usually booked in at short notice. We have to give good service to our (minor and major) servicing customers, but there is some flexibility in planning these jobs. At the other extreme, short-term emergency repair work for customers has to be fitted into our schedule as quickly as possible. If someone is desperate to have their car repaired at very short notice, we sometimes ask them to drop their car in as early as they can and pick it up as late as possible. This gives us the maximum amount of time to fit it into the schedule. There are a number of service options open to customers. We can book in short jobs for a fixed time and do it while they wait. Most commonly, we ask the customer to leave the car with us and collect it later.'

The technicians

The company employed five technicians, two trainee technicians, three part-time valeting assistants (who cleaned customers' cars), two part-time receptionists, Diya and two office assistants who reported to her. Each technician worked in their own service bay, with the trainees assisting them as required. Two of the technicians had worked for Dan since the company was founded. They were the most experienced and tended to be allocated a mixture of major servicing jobs and 'short-term' repairs that might require more advanced diagnostic work. One of the other technicians was only recently qualified and was usually allocated the more routine jobs, such as minor servicing and MOT checks (the UK Government requires vehicles over a certain age to be tested every year, these are known as 'Ministry of Transport' or MOT tests). The remaining two technicians were allocated a mixture of work. '*We are going to have to reconsider how we allocate jobs in the near future. The more experienced technicians will always do more of the repair and major servicing than the junior technicians, but we can't keep on giving them all the interesting work. The junior people will get frustrated if they only do routine work. Also, we have tended to keep the senior technicians lightly loaded so that they can be free for the unpredictable short-term repairs. This means that, if demand is lighter than usual, they are less heavily loaded than the others, However, they are quite good at helping the others out, if this happens*.' (Diya Chopra)

Scheduling the service bays

Most days, the service centre had to deal with 15–30 jobs, taking from half-an-hour up to a whole day, or very occasionally even longer. Most jobs had a time allowance. A minor service was usually allowed one-and-a-half hours, a major service around twice as long, and MOT tests would take half an hour. Short-term repairs could take anything between half an hour and all day. Diya would make an estimate of how long the job would take, often in consultation with one of the senior technicians, but some jobs would be difficult to estimate. '*Some jobs are easier than others to estimate. One of the guys will say something like, "Sounds like a cam belt, should take a couple of hours". At other times they will say, "Goodness knows, (or words to that effect), sorry, don't know how long it will take". That is why I have to leave time free in the week's schedule*.'

Figure 10.14 shows a typical schedule for the service centre at the beginning of the week, before many 'short-term' jobs have been allocated. At this stage, the schedule is purely nominal. For example, this particular week, only one 'short-term' job had been programmed into the schedule by early Monday morning, but within a couple of hours the unused space on the schedule would fill up. Often within an hour of starting work on a Monday morning, the schedule would have changed. Early in the week, this was usually because a standard service had taken longer than expected because a problem had been found. As Diya explained: '*Every day we have to cope with the unexpected. A technician may find that extra work is needed, customers may want extra work doing, and technicians are sometimes ill, which reduces our capacity. Occasionally parts may not be available so we have to arrange with the customer for the vehicle to be rebooked for a later time. Every week, up to ten or twelve customers just don't turn up. We automatically text them the day before, but even so they still forget to bring their car in so we have to rebook them in at a later time. We can cope with most of these uncertainties because our technicians can be flexible and most are also willing to work overtime when needed. The important thing is to manage customers' expectations. If there is a chance that the vehicle may not be ready for them, it shouldn't come as a surprise when they try to collect it.*'

Even with some flexibility from the technicians, as the week progressed short-term repairs were increasingly likely to disrupt the schedule. The actual schedules were recorded on a computer-based scheduling system that the centre had been using for several years. Diya found the system useful, but limited. '*We enter all jobs into the scheduling system. On screen it shows the total capacity we have day-by-day, all the jobs that are booked in, the amount of free capacity still available, and so on. We use this to see when we have the capacity to book a customer in, and then enter all the customer's details. The car maker has issued "standard times" for all the major jobs. However, you have to modify these standard times a bit to take account of circumstances. That is where the senior technicians' experience comes in. Of course, the system does not really make any decisions as to which jobs have priority, nor can it automatically reschedule when things change. We make all the decisions. It's really a convenient system for keeping track of what's happening. It also works out overtime payments, issues invoices and calculates workshop utilisation figures*.'

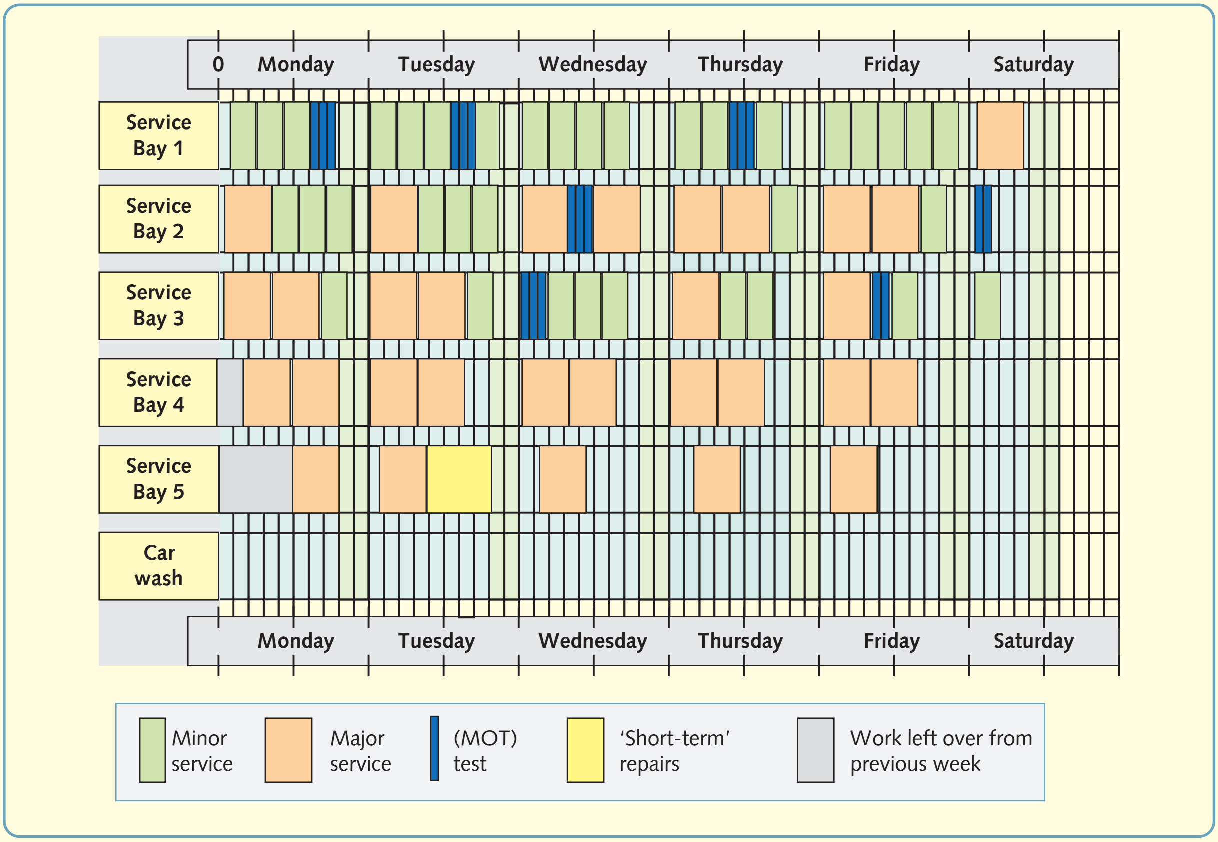


Figure 10.14 **Audall Auto Servicing schedule for upcoming week (as of Monday 08.00)**

The car wash and valeting

An unexpected bottleneck in the workshop was the valeting (including the car wash) facility. Diya explained: *'We don't valet all vehicles. Normally we process around 50–60 vehicles a day; valeting each of them takes between 5 and 15 minutes. There should be plenty of capacity, but very often vehicles are waiting for up to an hour before they can be cleaned. This means that customers can be waiting for longer than they expect, and longer than we have promised them. Dan firmly believes that we should never return a car (with the exception of minor repairs and MOT checks) that has not been fully valeted, but I know that, under pressure, the valeting staff sometimes rush the cleaning. I don't think customers mind too much.'*

Spare parts

Parts stock could be an important factor in keeping to schedule. Diya thought that their spare parts stock control was pretty good, but could be better: *'We keep all the most commonly used parts in stock, but if a repair needs a part which is not in stock, we can usually get it from our parts distributors within a day. Every evening, our planning system prints out the jobs that we should be doing the next day and the parts that are likely to be needed for each job. This allows the parts staff to pick out the parts for each job so that the technicians can receive them first thing the next morning without any delay. The problem is that because we normally get a part within a day, we can become complacent. If a part is not available for any reason it can delay a repair, which disrupts our schedule and can upset the customer. Maybe we should keep some "emergency stocks" of parts that we have had problems getting quickly, even if they are not often needed. It would tie up cash, but would help to protect the schedule.'*

Direct customer reservations

Dan and Diya had been investigating the costs and benefits of investing in a new computer-based scheduling system. There were several systems on the market, specifically designed for scheduling automobile workshops. Many of them were intended for larger workshops than Audall's, but some incorporated a new interface that would allow customers to book their vehicle into the workshop in a specific time slot. Dan was intrigued by the possibility of being able to do this, but was also hesitant. *'It would mean a very different way of working. I'm not sure whether it would be worth doing, especially in the way it would impact on customer service. For example, at the moment Diya can judge the degree of urgency of each job and talk to the customer, managing their expectations where appropriate.*

She is particularly good at making decisions that reconcile the needs of the customers and our resources in some way. For Diya, this involves attempting to maximise the utilisation of her workshop resources while keeping customers satisfied.'

Questions

1. The schedule in Figure 10.14 indicates that (a) the more predictable jobs tend to be loaded earlier in the day, and (b) technicians are allocated different types of job. If these two assumptions are true, do they seem sensible?
2. Is Dan right to be concerned that the workshop's utilisation is only 83 per cent?
3. How should Diya make a decision about keeping 'emergency stocks' of parts for which delivery times are uncertain?
4. Why is the valeting facility such a bottleneck when there should be plenty of capacity?
5. What do you see as the advantages and disadvantages of allowing direct customer reservations?

Applying the principles

Some of these exercises can be answered by reading the chapter. Others will require some general knowledge of business activity and some might require an element of investigation. **Model answers for the first two questions are available on the student companion website.** *For model answers to the other questions in this section, please ask your tutor/lecturer.*

1. Mark Key is an events coordinator for a small company. Returning from his annual holiday in France, he is given six events to plan in September. He gives them the codes A–F. He needs to decide upon the sequence in which to plan the events and wants to minimise the average time the jobs are tied up in the office and, if possible, meet the deadlines allocated. The six jobs are detailed in Table 10.1.

Table 10.1 The six jobs that Mark has to sequence

Coded jobs	*Process time (days)*	*Due date Sept*
A	4	12
B	3	5
C	1	7
D	2	9
E	2	15
F	5	8

Determine a sequence based on using:

a) the FIFO rule;
b) the due date rule; and
c) the shortest operation time rule.

Which of these sequences gives the most efficient solution and which gives the least lateness?

2. It is week 35 of a busy year at Ashby Architects and Jo Ashby is facing a big problem. Both of her junior partners have been diagnosed with a serious illness contracted on a trip to scope out a prospective job in Lichtenstein. So Jo has to step in and complete the outstanding jobs that were being worked on by the two juniors. The outstanding jobs are shown in Table 10.2.

Table 10.2 Outstanding jobs that Jo will need to complete

Job description	*Due date (week of 52)*	*Weeks of work remaining*
Ashthorpe lavatory block	40	2.0
Bugwitch bus shelters	48	5.0
Crudstone plc HQ	51	3.0
Dredge sewage works	52	8.0

Jo has heard that a sequencing rule called the critical ratio (CR) will give efficient results. The priority of jobs using the CR rule is defined by an index computed as follows:

$$CR = \frac{\text{time remaining}}{\text{work days remaining}} = \frac{\text{due date} - \text{today's date}}{\text{work days remaining}}$$

Using this rule, what priority should Jo give the jobs?

3. It takes six hours for a contract laundry to wash, dry and press (in that order) a batch of overalls. It takes 3 hours to wash the batch, 2 hours to dry it, and 1 hour to press it. Usually each day's batch is collected and ready for processing at 8.00 a.m. and needs to be picked up at 4.00 p.m. The two people who work in the laundry have different approaches to how they schedule the work. One schedules 'forward'. Forward scheduling involves starting work as soon as it arrives. The other schedules 'backward'. Backward scheduling involves starting jobs at the last possible moment that will prevent them from being late.

a) Draw up a schedule indicating the start and finish time for each activity (wash, dry and press) for both forward and backward approaches.

b) What do you think are the advantages and disadvantages of these two approaches?

4. Read the following descriptions of two cinemas.

Kinepolis in Brussels is one of the largest cinema complexes in the world, with 28 screens, a total of 8,000 seats, and four showings of each film every day. It is equipped with the latest projection technology. All the film performances are scheduled to start at the same times every day: 4.00 p.m., 6.00 p.m. 8.00 p.m. and 10.30 p.m. Most customers arrive in the 30 minutes before the start of the film. Each of the 18 ticket desks has a networked terminal and a ticket printer. For each customer, a screen code is entered to identify and confirm seat availability for the requested film. Then the number of seats required is entered and the tickets are printed, though these do not allocate specific seat positions. The operator then takes payment by cash or credit card and issues the tickets. This takes an average of 19.5 seconds, and a further 5 seconds is needed for the next customer to move forward. An average transaction involves the sale of approximately 1.7 tickets.

The UCI cinema in Birmingham has eight screens. The cinema incorporates many 'state-of-the-art' features, including a high-quality THX sound system, fully computerised ticketing and a video games arcade off the main hall. In total the eight screens can seat 1,840 people; the capacity (seating) of each screen varies, so the cinema management can allocate the more popular films to the larger screens and use the smaller screens for the less popular films. The starting times of the eight films at UCI are usually staggered by 10 minutes, with the most popular film in each category (children's, drama, comedy, etc.) being scheduled to run first. Because the films are of different durations, and since the manager must try to maximise the utilisation of the seating, the scheduling task is complex. Ticket staff are continually aware of the remaining capacity of each 'screen' through their terminals. There are up to four ticket desks open at any one time. The target time per overall transaction is 20 seconds. The average number of ticket sales per transaction is 1.8. All tickets indicate specific seat positions, and these are allocated on a first come, first served basis.

a) What are the main differences between the two cinemas from the perspectives of their operations managers?

b) What are the advantages and disadvantages of the two different methods of scheduling the films onto the screens?

c) Find out the running times and classifications of eight popular films. Try to schedule these onto the UCI Birmingham screens, taking account of what popularity you might expect at different times. Allow 20 minutes for emptying, cleaning and admitting the next audience, and 15 minutes for advertising before the start of each film.

5. Think through the following three brief examples. What type of control (according to Figure 10.12) do you think each one warrants?

a) The Games Delivery Authority (GDA) was a public body responsible for developing and building the new venues and infrastructure for the 'International Games' and their use after the event. The GDA appointed a consortium responsible for the overall programme's quality, delivery and cost, in addition to health and safety, sustainability, equality and diversity targets. The Games Park was a large construction programme spreading across five separate local government areas, including transport developments, retail areas and local regeneration projects. Sustainability was central to the development. '*Sustainability' was ingrained into our thinking – from the way we planned, built and worked, to the way we play, socialise and travel*.' To ensure they stuck to commitments, the GDA set up an independent body to monitor the project. All potential contractors tendering for parts of the project were aware that a major underlying objective of the Games initiative was regeneration. The Games site was to be built on highly industrialised and contaminated land.

b) The supermarket's new logistics boss was blunt in his assessment of its radical supply chain implementation: '*Our rivals have watched in utter disbelief*', he said. '*Competitors looked on in amazement as we poured millions into implementing new IT systems and replaced 21 depots with a handful of giant automated "fulfilment factories"*.' '*In hindsight, the heavy reliance on automation was a big mistake, especially for fast-moving goods*', said the company's CIO. '*When a conventional facility goes wrong, you have lots of options. You have flexibility to deal with issues. When an automated "fulfilment factory" goes wrong, frankly, you're stuck*.' Most damning was the way that the supermarket pressed on with the implementation of the automated facilities before proving that the concept worked at the first major site. '*I'd have at least proved that one of them worked before building the other three*', the CIO said. '*Basically, the whole company was committed to doing too much, too fast, trying to implement a seven-year strategy in a three-year timescale*.'

c) '*It's impossible to overemphasise just how important this launch is to our future*', said the CEO. '*We have been losing market share for seven quarters straight. However, we have very high hopes for the new XC10 unit*.' And most of the firm's top management team agreed with her. Clearly the market had been maturing for some time now, and was undoubtedly getting more difficult. New product launches from competitors had been eroding both market share and margin. Yet competitors' products, at best, simply matched the firm's offerings in all benchmark tests. '*Unless someone comes up with a totally new technology, which is very unlikely, it will be a matter of making marginal improvements in product performance and combining this with well-targeted and coordinated marketing. Fortunately, we are good at both of these. We know this technology, and we know these markets. We are also clear what role the new XC10 should play. It needs to consolidate our market position as the leader in this field, half the slide in market share, and re-establish our customers' faith in us. Margins, at least in the short term, are less important*.'

Notes on chapter

1 The information on which this example is based is taken from: Masresha, A.K. (2011) *Aircraft Scheduling: Basics, models and applications at Ethiopian Airlines*, Lambert Academic Publishing; Farman, J. (1999) 'Les Coulisses du Vol', Air France. Talk presented by Richard E. Stone, NorthWest Airlines at the IMA Industrial Problems Seminar, 1998.

2 Wight, O. (1984) *Manufacturing Resource Planning MRP II: Unlocking America's productivity challenge*, Oliver Wight Publications.

3 The information on which this example is based is taken from: Fruhlinger, J., Wailgum, T. and Sayer, P. (2020) '16 famous ERP disasters, dustups and disappointments', CIO. com, 20 March, https://www.cio.com/article/2429865/enterprise-resource-planning-10-famous-erp-disasters-dustups-and-disappointments.html [accessed 25 September 2020]; Novacura (2019) '4 ERP implementation failures with valuable lessons', The Novacura Flow blog, 19 February, https://www2.novacura.com/blog/why-do-erp-implementations-fail [accessed 25 September 2020]; Kanaracus, C. (2008) 'Waste Management sues SAP over ERP implementation', InfoWorld, 27 March; https://www.lidl.co.uk/en; https://www.woolworths.com.au/; https://www.oriola.com/; https://www.wm.com/ [all accessed 7 October 2020].

4 Coldrick, A., Ling, D. and Turner, C. (2003) 'Evolution of sales and operations planning – from production planning to integrated decision making', StrataBridge Working Paper.

5 The information on which this example is based is taken from: Economist (2020) 'Triage under trial: The tough ethical decisions doctors face with covid-19', *Economist* print edition, 2 April; Jones, C. (2020) 'What a career in intensive care nursing has taught me about triage', *Financial Times*, 6 April.

6 Economist (2011) 'Please be seated: A faster way of boarding planes could save time and money', *Economist* print edition, 3 September; Palmer, J. (2011) 'Tests show fastest way to board passenger planes', BBC News website, 31 August.

7 The information on which this example is based is taken from: Heathrow website, https://www.heathrow.com; for a technical explanation of the aircraft landing algorithm, see Beasley, J.E., Sonander, J. and Havelock, P. (2001) 'Scheduling aircraft landings at London Heathrow using a population heuristic', *Journal of the Operational Research Society*, 52 (5), pp. 483–493.

8 Goldratt, E.Y. and Cox, J. (2004) *The Goal: A process of ongoing improvement*, 3rd edition, Gower Publishing Ltd.

9 This approach is based on an original idea described in Hofstede, G. (198l) 'Management control of public and not-for-profit activities', *Accounting, Organizations and Society*, (6) 3, pp. 193–211.

Taking it further

Akhtar, J. ***(2016) Production Planning and Control with SAP ERP, 2nd edition, SAP Press.*** *Comprehensive, but specialised and focused on one particular ERP vendor (SAP).*

Berry, W., Vollmann, T.E., Jacobs, T.F.R. and Whybark, D.C. ***(2011) Manufacturing Planning and Control System for Supply Chain Management, McGraw-Hill Education.*** *This is the bible of production planning and control. It deals with all the issues covered in this chapter.*

Bradford, M. ***(2015) Modern ERP: Select, implement, and use today's advanced business systems, LuLu.com.*** *Very much for those wanting detailed ERP information.*

Chapman, S.N. ***(2005) Fundamentals of Production Planning and Control, Pearson.*** *A good basic textbook.*

Goldratt, E.Y. and Cox, J. ***(2004) The Goal: A process of ongoing improvement, 3rd edition, Gower Publishing Ltd.*** *Don't read this if you like good novels but do read this if you want an enjoyable way of understanding some of the complexities of scheduling. It particularly applies to the drum, buffer, rope concept described in this chapter and it also sets the scene for the discussion of OPT.*

Magal, S.R. and Word, J. ***(2010) Integrated Business Processes with ERP Systems, John Wiley & Sons.*** *It's written in partnership with SAP, the leading seller of ERP systems, but it does cover all of the key processes supported by modern ERP systems.*

Supplement to Chapter 10

Materials requirements planning (MRP)

Materials requirements planning (MRP) is an approach to calculating how many parts or materials of particular types are required and what times they are required. To do this, data files are needed which, when the MRP program is run, can be checked and updated. Figure 10.15 shows how these files relate to each other. The first inputs to materials requirements planning are customer orders and forecast demand. MRP performs its calculations based on the combination of these two parts of future demand. All other requirements are derived from, and dependent on, this demand information.

Master production schedule

The master production schedule (MPS) forms the main input to materials requirements planning and contains a statement of the volume and timing of the end products to be made. It drives all the production and supply activities that will eventually come together to form the end products. It is the basis for the planning and utilisation of labour and equipment, and it determines the provisioning of materials and cash. The MPS should include all sources of demand, such as spare parts, internal production promises, and so on. For example, if a manufacturer of earth excavators plans an exhibition of its products and allows a project team to raid the stores so that it can build two pristine examples to be exhibited, this is likely to leave the factory short of parts. MPS can also be used in service organisations. For example, in a hospital operating theatre there is a master schedule that contains a statement of which operations are planned and when. This can be used to provision materials for the operations, such as the sterile instruments, blood and dressings. It may also govern the scheduling of staff for operations.

The master production schedule record

Master production schedules are time-phased records of each end product, which contain a statement of demand and currently available stock of each finished item. Using this information, the available inventory is projected ahead in time. When there is insufficient inventory to satisfy forward demand, order quantities are entered on the master schedule line. Table 10.3 is a simplified example of part of a master production schedule for one item. In the first row

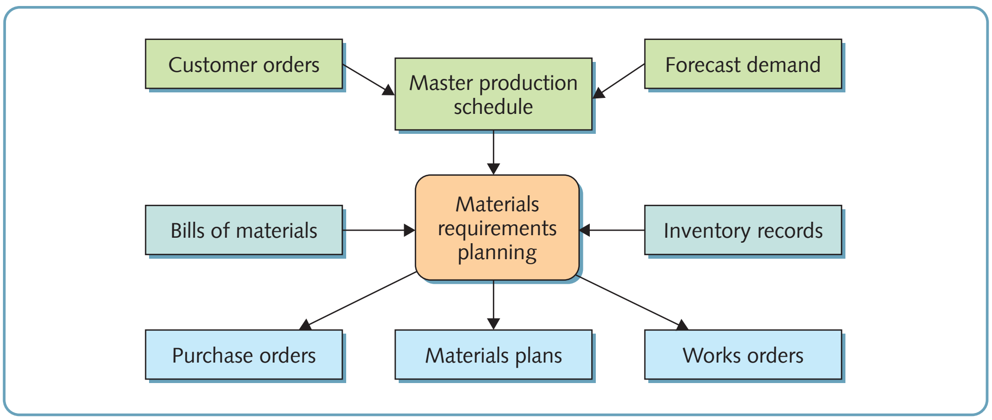


Figure 10.15 **Materials requirements planning (MRP) schematic**

Table 10.3 Example of a simple master production schedule

		Week number								
		1	2	3	4	5	6	7	8	9
Demand		10	10	10	10	15	15	15	20	20
Available		20	10	0	0	0	0	0	0	0
MPS		0	0	10	10	15	15	15	20	20
On hand	30									

Table 10.4 Example of a 'level' master production schedule

		Week number								
		1	2	3	4	5	6	7	8	9
Demand		10	10	10	10	15	15	15	20	20
Available		31	32	33	34	30	26	22	13	4
MPS		11	11	11	11	11	11	11	11	11
On hand	30									

Table 10.5 Example of a level master production schedule including available to promise

		Week number								
		1	2	3	4	5	6	7	8	9
Demand		10	10	10	10	15	15	15	20	20
Sales orders		10	10	10	8	4				
Available		31	32	33	34	30	26	22	13	4
ATP		31	1	1	3	7	11	11	11	11
MPS		11	11	11	11	11	11	11	11	11
On hand	30									

the known sales orders and any forecast are combined to form 'Demand'. The second row, 'Available', shows how much inventory of this item is expected to be in stock at the end of each weekly period. The opening inventory balance, 'On hand', is shown separately at the bottom of the record. The third row is the master production schedule, or MPS; this shows how many finished items need to be completed and available in each week to satisfy demand.

Chase or level master production schedules

In the example in Table 10.3, the MPS increases as demand increases and aims to keep available inventory at 0. The master production schedule is 'chasing' demand (see Chapter 8) and so adjusting the provision of resources. An alternative 'levelled' MPS for this situation is shown in Table 10.4. Level scheduling involves averaging the amount required to be completed to smooth out peaks and troughs; it generates more inventory than the previous MPS.

Available to promise (ATP)

The master production schedule provides the information to the sales function on what can be promised to customers and when delivery can be promised. The sales function can load known sales orders against the master production schedule and keep track of what is available to promise (ATP) (*see* Table 10.5). The ATP line in the master production schedule shows the maximum that is still available in any one week, against which sales orders can be loaded.

The bill of materials (BOM)

From the master schedule, MRP calculates the required volume and timing of assemblies, sub-assemblies and materials. To do this it needs information on what parts are required for each product. This is called the 'bill of materials'. Initially it is simplest to think about this as a product structure. The product structure in Figure 10.16 is a simplified structure showing the parts required to make a simple board game. Different 'levels of assembly' are shown with the finished product (the boxed game) at level 0, the parts and sub-assemblies that go into the boxed game at level 1, the parts that go into the sub-assemblies at level 2, and so on.

A more convenient form of the product structure is the 'indented bill of materials'. Table 10.6 shows the whole indented bill of materials for the board game. The term 'indented' refers to the indentation of the level of assembly, shown in the left-hand column. Multiples of some parts are required; this means that MRP has to know the required number of each part to be able to multiply up the requirements. Also, the same part (for example, the TV label, part number 10062) may be used in different parts of the product structure. This means that MRP has to cope with this commonality of parts and, at some stage, aggregate the requirements to check how many labels in total are required.

Inventory records

MRP calculations need to recognise that some required items may already be in stock. So it is necessary, starting at level 0 of each bill, to check how much inventory is available of each finished product, sub-assembly and component, and then to calculate what is termed the 'net' requirements – that is, the extra requirements needed to supplement the inventory so that demand can be met. This requires that three main inventory records are kept: the item master file, which contains the unique standard identification code for each part or component; the transaction file, which keeps a record of receipts into stock, issues from stock and a running balance; and the location file, which identifies where inventory is located.

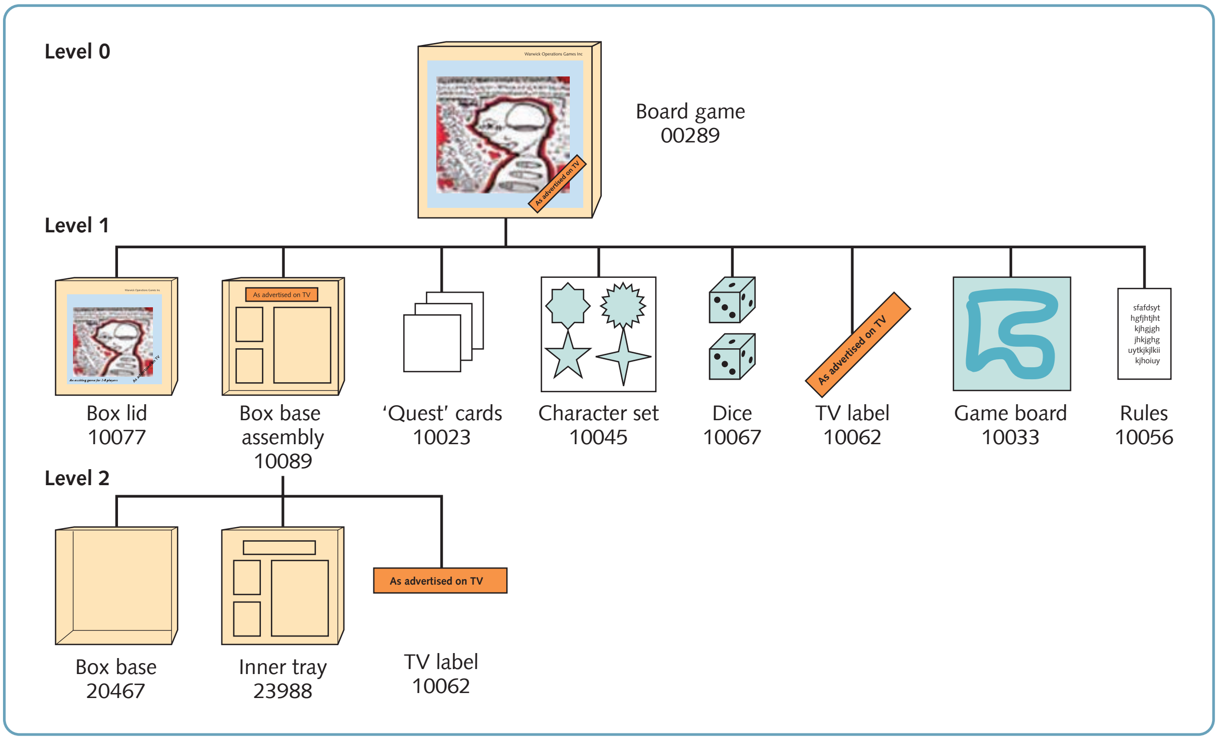


Figure 10.16 **Product structure for a simple board game**

Table 10.6 Indented bill of materials for board game

Part number: *00289* Description: *Board game* Level: *0*			
Level	*Part number*	*Description*	*Quantity*
0	00289	Board game	1
.1	10077	Box lid	1
.1	10089	Box base ass.	1
. .2	20467	Box base	1
. .2	10062	TV label	1
. .2	23988	Inner tray	1
.1	10023	'Quest' cards set	1
.1	10045	Character set	1
.1	10067	Dice	2
.1	10062	TV label	1
.1	10033	Game board	1
.1	10056	Rules booklet	1

The MRP netting process

The information needs of MRP are important, but they are not the 'heart' of the MRP procedure. At its core, MRP is a systematic process of taking this planning information and calculating the volume and timing requirements that will satisfy demand. The most important element of this is the MRP netting process.

Figure 10.17 illustrates the process that MRP performs to calculate the volumes of materials required. The master production schedule is 'exploded', examining the implications of the schedule through the bill of materials, checking how many sub-assemblies and parts are required. Before moving down the bill of materials to the next level, MRP checks how many of the required parts are already available in stock. It then generates 'works orders', or requests, for the net requirements of items. These form the schedule, which is again exploded through the bill of materials at the next level down. This process continues until the bottom level of the bill of materials is reached.

Back-scheduling

In addition to calculating the volume of materials required, MRP also considers when each of these parts is required – that is, the timing and scheduling of materials. It does this by a process called back-scheduling, which takes into account the lead time (the time allowed for completion of each stage of the process) at every level of assembly. Again, using the example of the board game, assume that 10 board games are required to be finished by a notional planning day, which we will term day 20. To determine when we need to start work on all the parts that make up the game, we need to know all the lead times that are stored in MRP files for each part (see Table 10.7).

Using the lead-time information, the program is worked backwards to determine the tasks that have to be performed and the purchase orders that have to be placed. Given the lead times and inventory levels shown in Table 10.5, the MRP records shown in Figure 10.18 can be derived.

MRP capacity checks

The MRP process needs a feedback loop to check whether a plan was achievable and whether it has actually been achieved. Closing this planning loop in MRP systems involves checking production plans against available capacity and, if the proposed plans are not achievable at any level, revising them. All but the simplest MRP systems are now closed-loop systems. They use three planning routines to check production plans against the operation's resources at three levels:

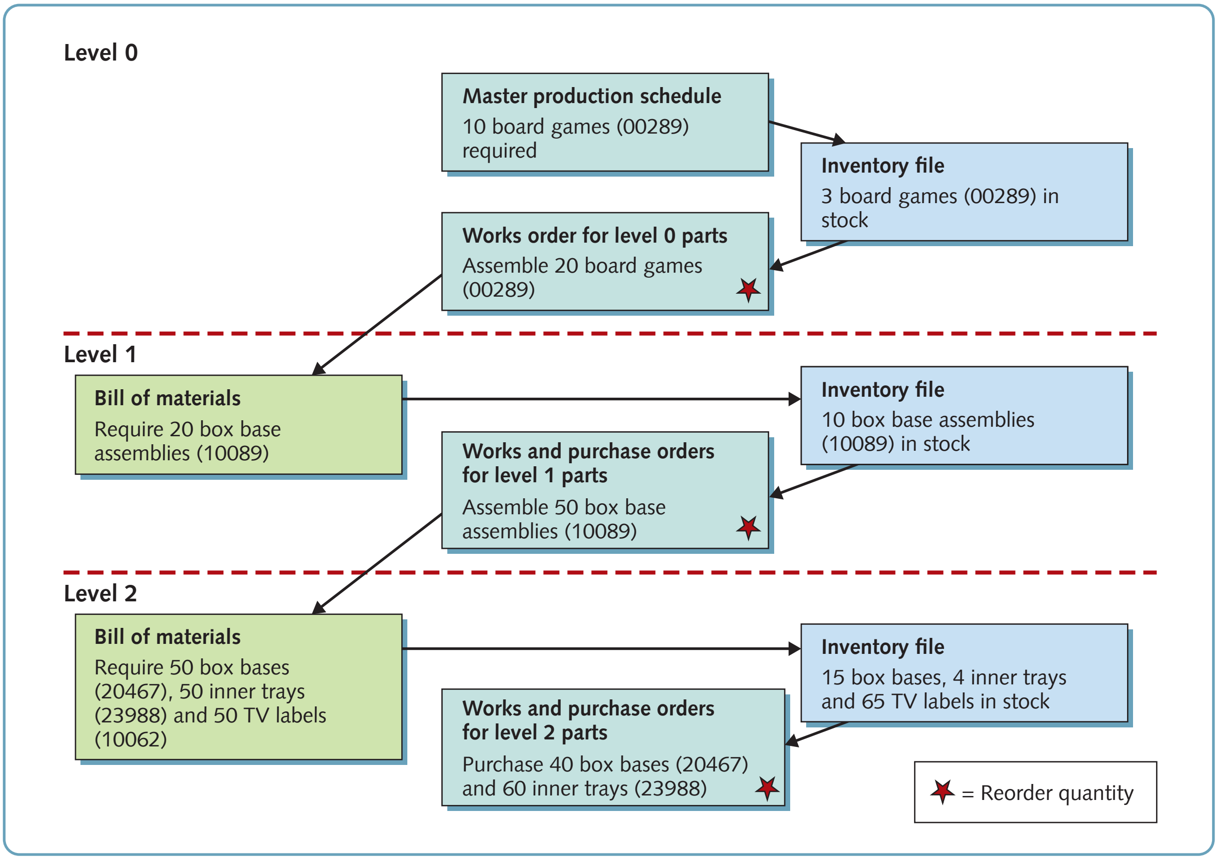


Figure 10.17 The MRP netting calculations for the simple board game

Table 10.7 Back-scheduling of requirements in MRP

Part no.	*Description*	*Inventory on hand day 0*	*Lead time (days)*	*Reorder quantity*
00289	Board game	3	2	20
10077	Box lid	4	8	25
10089	Box base ass.	10	4	50
20467	Box base	15	12	40
23988	Inner tray	4	14	60
10062	TV label	65	8	100
10023	'Quest' cards set	4	3	50
10045	Character set	46	3	50
10067	Dice	22	5	80
10033	Game board	8	15	50
10056	Rules booklet	0	3	80

1. ***Resource requirements plans (RRPs)*** – involve looking forward in the long term to predict the requirements for large structural parts of the operation, such as the numbers, locations and sizes of new plants.
2. ***Rough-cut capacity plans (RCCPs)*** – are used in the medium to short term, to check the master production schedules against known capacity bottlenecks, in case capacity constraints are broken. The feedback loop at this level checks the MPS and key resources only.
3. ***Capacity requirements plans (CRPs)*** – look at the day-to-day effect of the works orders issued from the MRP on the loading individual process stages.

00289: Treasure hunt game Assembly lead time = 2 Reorder quantity = 20

Day number:	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Requirements gross																					10
Scheduled receipts																					
On hand inventory	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	13
Planned order release																			20		

10077: Box lid Purchase lead time = 8 Reorder quantity = 25

Day number:	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Requirements gross																			20		
Scheduled receipts																					
On hand inventory	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	9	9	9
Planned order release											25										

10089: Box base assembly Assembly lead time = 4 Reorder quantity = 50

Day number:	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Requirements gross																			20		
Scheduled receipts																					
On hand inventory	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	40	40	40
Planned order release															50						

20467: Box base Purchase lead time = 12 Reorder quantity = 40

Day number:	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Requirements gross															50						
Scheduled receipts																					
On hand inventory	15	15	15	15	15	15	15	15	15	15	15	15	15	15	5	5	5	5	5	5	5
Planned order release			40																		

23988: Inner tray Purchase lead time = 14 Reorder quantity = 60

Day number:	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Requirements gross															50						
Scheduled receipts																					
On hand inventory	4	4	4	4	4	4	4	4	4	4	4	4	4	4	14	14	14	14	14	14	14
Planned order release	60																				

10062: TV label Purchase lead time = 8 Reorder quantity = 100

Day number:	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Requirements gross															50				20		
Scheduled receipts																					
On hand inventory	65	65	65	65	65	65	65	65	65	65	65	65	65	65	15	15	15	15	95	95	95
Planned order release											100										

10023: Quest card set Purchase lead time = 3 Reorder quantity = 50

Day number:	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Requirements gross																			20		
Scheduled receipts																					
On hand inventory	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	34	34	34
Planned order release																50					

10045: Character set Purchase lead time = 3 Reorder quantity = 50

Day number:	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Requirements gross																			20		
Scheduled receipts																					
On hand inventory	46	46	46	46	46	46	46	46	46	46	46	46	46	46	46	46	46	46	26	26	26
Planned order release																					

10067: Dice Purchase lead time = 5 Reorder quantity = 80

Day number:	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Requirements gross																			40		
Scheduled receipts																					
On hand inventory	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	13
Planned order release														80							

10033: Game board Purchase lead time = 15 Reorder quantity = 50

Day number:	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Requirements gross																			20		
Scheduled receipts																					
On hand inventory	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	38	38	38
Planned order release				50																	

10056: Rules booklet Purchase lead time = 3 Reorder quantity = 80

Day number:	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Requirements gross																			20		
Scheduled receipts																					
On hand inventory	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	60	60	60
Planned order release																80					

Figure 10.18 Extract from the MRP records for the simple board game (lead times indicated by arrows →←)

11 Lean synchronisation

Introduction

Lean synchronisation uses 'lean' (what was once called 'just-in-time' or 'JIT') principles to supply perfect-quality products and services in synchronisation with the demand for them, with zero waste and at low cost. Once a radical departure from traditional operations practice, these principles have transformed the way leading organisations think about their operations. Although the topic is sometimes narrowly treated as a manufacturing phenomenon (unsurprisingly, given the pioneering role of Toyota in lean management), lean synchronisation principles can be applied across all sectors, including finance, healthcare, IT, retailing, construction, agriculture and the public sector. Figure 11.1 shows the position of the ideas described within this chapter in the general model of operations management.

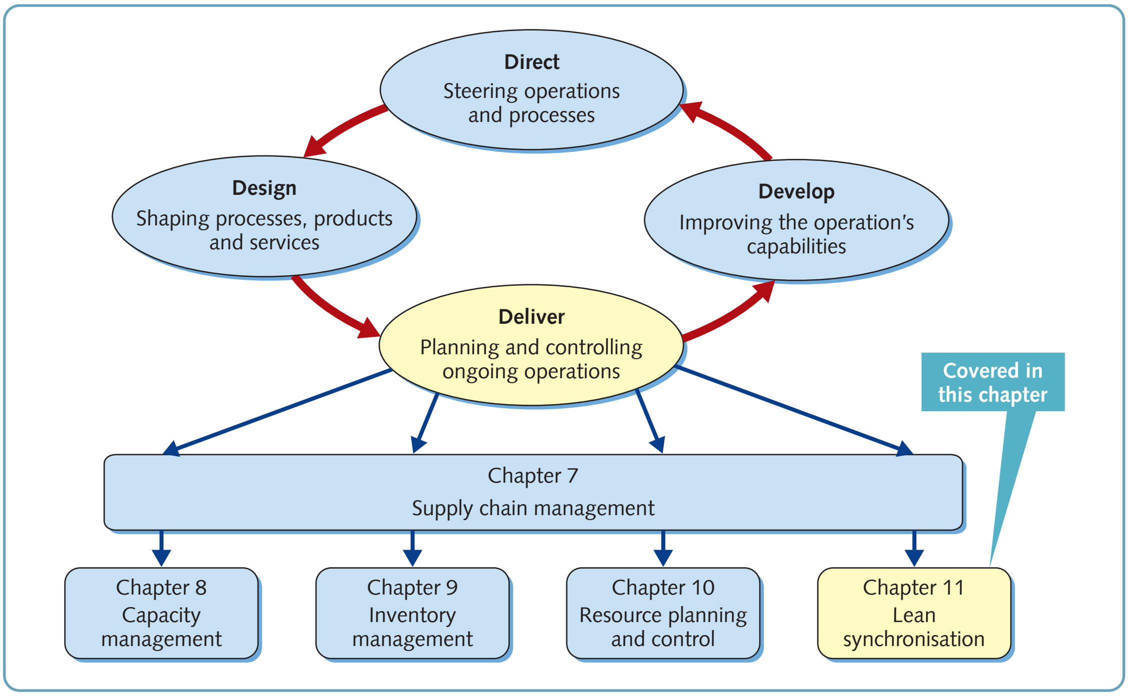
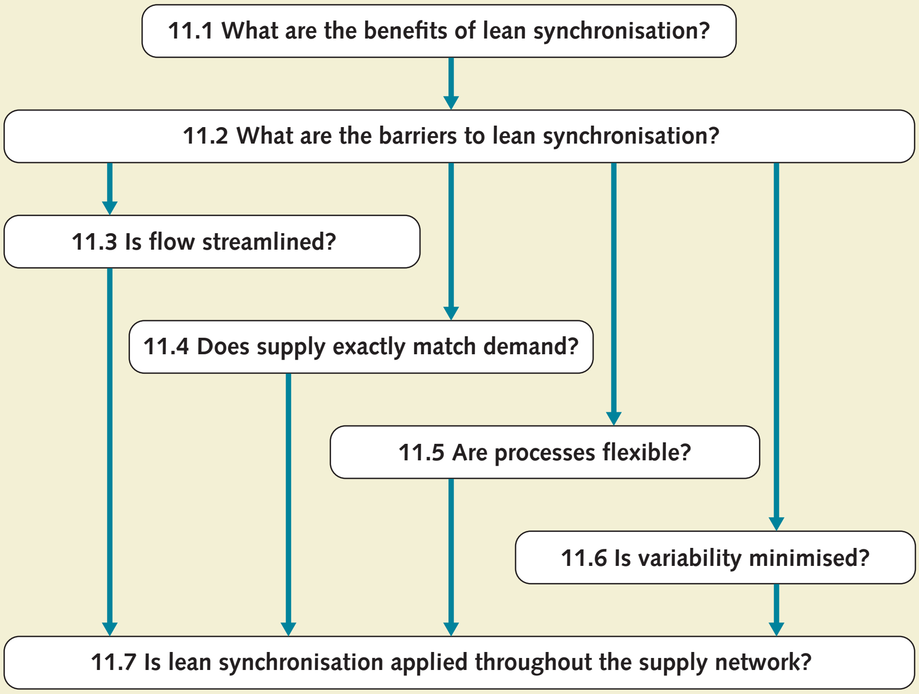


Figure 11.1 Lean synchronisation has the aim of achieving a flow of products and services that always delivers exactly what customers want, in exact quantities, exactly when needed, exactly where required and at the lowest possible cost

EXECUTIVE SUMMARY



11.1 What are the benefits of lean synchronisation?

The focus of lean synchronisation is to achieve a flow of products and services that delivers exactly what customers want, in exact quantities, exactly when needed, exactly where required and at the lowest possible cost. It is a concept that is almost synonymous with terms such as 'just-in-time' (JIT) and 'lean operations principles'. The central benefit is that if products, customers or information flow smoothly, not only is throughput time reduced, but the negative effects of in-process inventory or queues are avoided. Inventories and queues are seen as obscuring the problems that exist within processes and therefore inhibiting process improvement.

11.2 What are the barriers to lean synchronisation?

The aim of lean synchronisation can be inhibited in three ways. First is the failure to eliminate waste in all parts of the operation; and the causes of waste are more extensive than is generally understood. The second is a failure to involve all the people within the operation in the shared task of smoothing flow and eliminating waste; Japanese proponents of lean synchronisation often use a set of 'basic working practices' to ensure involvement. Third is the failure to adopt continuous improvement principles. Because pure lean synchronisation is an aim rather than something that can be implemented quickly, it requires the continual application of incremental improvement steps to reach its potential.

11.3 Is flow streamlined?

Long process routes are wasteful and cause delay and inventory build-up. Physically reconfiguring processes to reduce distance travelled and aid cooperation between staff can help to streamline flow. Similarly, ensuring flow visibility helps to make improvement to flow easier. Sometimes this can involve small-scale technologies that can reduce fluctuations in flow volume.

11.4 Does supply exactly match demand?

The aim of lean synchronisation is to meet demand exactly – neither too much nor too little, and only when it is needed. Pull control principles are typically used to achieve this goal. The most common method of doing this is the use of kanbans – simple signalling devices that prevent the accumulation of excess inventory.

11.5 Are processes flexible?

Responding exactly to demand only when it is needed often requires a degree of flexibility in processes, both to cope with unexpected demand and to allow processes to change between different activities without excessive delay. This often means reducing changeover times in technologies.

11.6 Is variability minimised?

Variability in processes disrupts flow and prevents lean synchronisation. Variability includes quality variability and schedule variability. Statistical process control (SPC) principles are useful in reducing quality variability. The use of levelled scheduling and mixed modelling can be used to reduce flow variability, and total productive maintenance (TPM) can reduce the variability caused by breakdowns.

11.7 Is lean synchronisation applied throughout the supply network?

The same benefits and principles of lean synchronisation that apply within operations can also apply across supply networks. This is more difficult, partly because of the complexity of flow and partly because supply networks are prone to the type of unexpected fluctuations that are easier to control within operations.

11.1 Diagnostic question: *What are the benefits of lean synchronisation?*

Synchronisation means that the flow of products and services always delivers exactly what customers want (perfect quality), in exact quantities (neither too much nor too little), exactly when needed (not too early nor too late) and exactly where required (not to the wrong location). Lean synchronisation is to do all this at the lowest possible cost. It results in customers, products and information flowing rapidly and smoothly through processes, operations and supply networks. As such, it is not an idea that is relevant only to one context. While many of the principles of lean described here emerged in manufacturing organisations initially, they are nearly all applicable to other sectors. In this chapter, we will provide examples of organisations in a wide range of sectors that apply facets of lean to their operations and supply networks. Many of the examples of lean philosophy and lean techniques in service industries are directly analogous to those found in manufacturing, because physical items are being moved or processed in some way. For example, supermarkets usually replenish their shelves only when customers have taken sufficient products off the shelf. The movement of goods from the 'back office' store to the shelf is triggered only by the 'empty shelf' demand signal. Construction companies increasingly make a rule of only calling for material deliveries to their sites the day before materials are needed. This reduces clutter and the chances of theft. Both are examples of the application of pull control principles. Other examples of lean concepts and methods apply even when most of the service elements are intangible. For example, new publishing technologies allow professors to assemble printed and e-learning course material customised to the needs of individual courses or even individual students. Here, we see the lean principles of flexibility and small batch sizes allowing customisation and rapid delivery.

The benefits of synchronised flow

The best way to understand how lean synchronisation differs from more traditional approaches to managing flow is to contrast the two simple processes in Figure 11.2. The traditional approach (a) assumes that each stage in the process will place its output in an inventory or queue that 'buffers' that stage from the next one downstream in the process. The next stage down will then (eventually) take outputs from this buffer, process them and pass them through to the next buffer inventory or queue. These buffers are there to 'insulate' each stage from its neighbours, making each stage relatively independent so that if, for example, stage A stops operating for some reason, stage B can continue, at least for a time. This insulation has to be paid for in terms of inventory or queues and slow throughput times because products, customers or information will spend time waiting between stages in the process.

The main argument against this traditional approach lies in the very conditions it seeks to promote, namely the insulation of the stages from one another. When a problem occurs at one stage, the problem will not immediately be apparent elsewhere in the system. The responsibility for solving the problem will be centred largely on the people within that stage, and the consequences of the problem will be prevented from spreading to the whole system. However, contrast this with the pure lean synchronised process illustrated in Figure 11.2 (b). Here products, customers or information are processed and then passed directly to the next stage 'just-in-time' for them to be processed further. Now if stage A stops processing, stage B will notice immediately and stage C very soon after. Stage A's problem is now quickly exposed to the whole process, which is immediately affected by the problem. This means that the responsibility for solving the problem is no longer confined to the staff at stage A. Everyone now shares the problem, considerably improving the chances of the problem being solved if only because

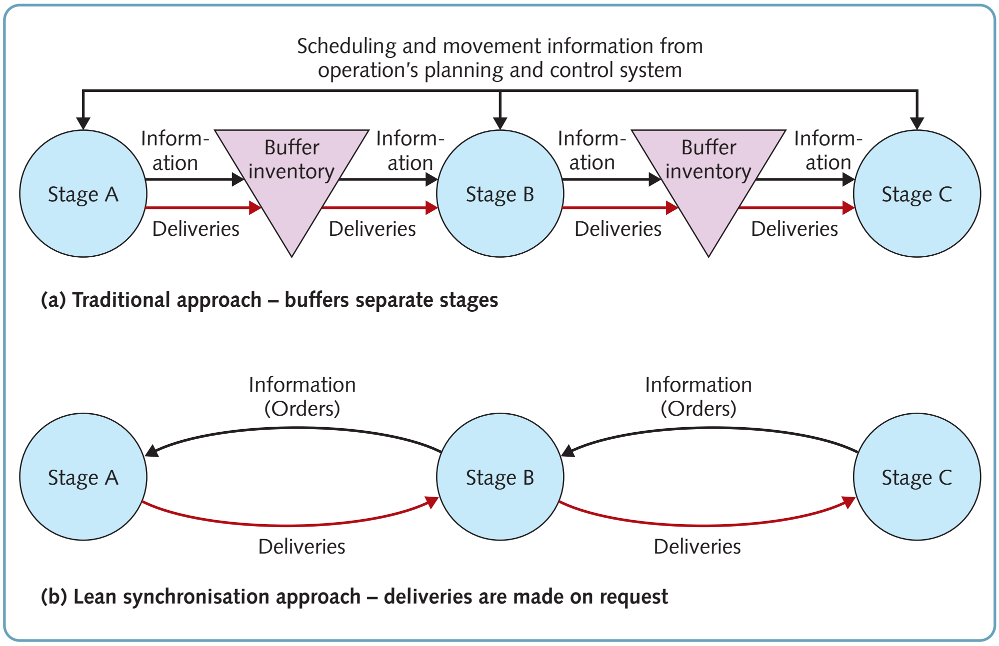


Figure 11.2 (a) Traditional and (b) lean synchronised flow between stages

it is now too important to be ignored. In other words, by preventing products, customers or information resources accumulating between stages, the operation has increased the chances of the efficiency of the operation being improved.

Non-synchronised approaches seek to encourage efficiency by protecting each part of the process from disruption. The lean synchronised approach takes the opposite view. Exposure of the system (although not suddenly, as in our simplified example) to problems can both make them more evident and change the motivation of the whole system towards solving the problems. Lean synchronisation sees accumulations of inventories, be they product, customer or information inventories, as a 'blanket of obscurity' that lies over the system and prevents problems being noticed. Table 11.1 shows the operational consequences of the build-up of these different inventory types.

Risks of low inventory

There has always been a risk of reducing inventories within processes, operations or supply networks, and that is that it increases vulnerability to large disruptions or changes in supply or demand. We deal with the more negative consequences of low inventory in the chapters on supply chain management (Chapter 7), inventory management (Chapter 9), and risk and resilience (Chapter 14). Obviously, any philosophy, as with lean, where inventory reduction is a central tenet is going to have to accept that its adoption inevitably incurs some degree of risk. The issue for any operations manager is whether that risk is worth it. In other words, do the benefits of increased exposure and sensitivity to relatively minor, yet cumulatively significant, disruptions outweigh the marginal reduction of resilience that lean processes incur? For some variation from normal operating conditions, for example a short-term and predictable decrease in supply, it would be sensible to allow inventories to (temporarily) increase. For more drastic disruptions, such as happened in the COVID-19 pandemic, for many industries both demand and supply were affected to such an extent that high or low inventory would have made little difference.

OPERATIONS PRINCIPLE
Buffer inventories (product, customer or information) used to insulate stages localise the motivation to improve.

Table 11.1 Operational implications of different inventory types

	Inventory		
	Material (queue of material)	***Information (queue of information)***	***Customers (queue of people)***
Cost	Ties up working capital	Less current information and so worth less	Wastes customers' time
Space	Needs storage space	Needs memory capacity	Need waiting area
Quality	Defects hidden, possible damage	Defects hidden, possible data corruption	Gives negative perception
Decoupling	Makes stages independent	Makes stages independent	Promotes job specialisation/ fragmentation
Utilisation	Stages kept busy by work in progress	Stages kept busy by work in data queues	Servers kept busy by waiting customers
Coordination	Avoids need for synchronisation	Avoids need for straight-through processing	Avoids having to match supply and demand

Source: Adapted from Fitzsimmons, J.A. (1990) 'Making continual improvement: A competitive strategy for service firms' in Bowen, D.E., Chase, R.B., Cummings, T.G. and Associates (eds), *Service Management Effectiveness*, Jossey-Bass.

The river and rocks analogy

The idea of obscuring the effects of inventory is often illustrated diagrammatically, as in Figure 11.3. The many problems of the operation are shown as rocks in a riverbed that cannot be seen because of the depth of the water. The water in this analogy represents the inventory in the operation. Yet, even though the rocks cannot be seen, they slow the progress of the river's flow and cause turbulence. Gradually reducing the depth of the water (inventory) exposes the worst of the problems that can be resolved, after which the water is lowered further, exposing more problems, and so on. The same argument applies for flow between whole processes or whole operations. For example, stages A, B and C in Figure 11.2 could be a supplier operation, a manufacturer and a customer's operation respectively.

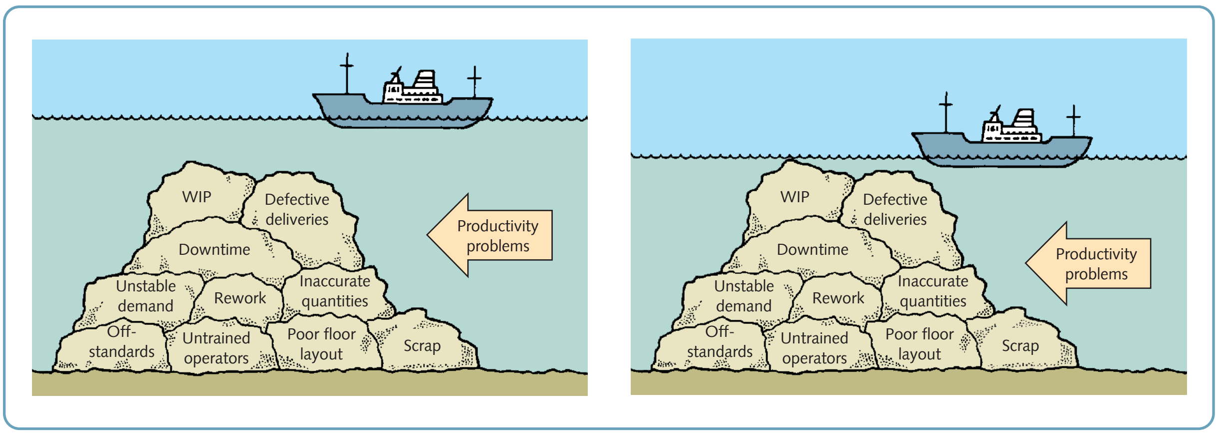


Figure 11.3 Reducing the level of inventory or queues (water) allows operations management (the ship) to see the problems in the operation (the rocks) and work to reduce them

Synchronisation, 'lean' and 'just-in-time'

Our definition of lean synchronisation is that it aims to meet demand instantaneously, with perfect quality and no waste. It is a concept that is almost synonymous with terms such as 'just-in-time' (JIT), the 'Toyota Production System' (TPS) and 'stockless production'. 'Just-in-time' emphasises the idea of producing items only when they are needed. 'Lean' stresses the elimination of waste. 'Stockless production' emphasises the negative effects of in-process inventory or queues. The 'Toyota Production System' encompasses all of these things. In fact, all these concepts overlap to a large degree, and no definition fully conveys the full implications for operations practice. Here we use the term lean synchronisation because it best describes the impact of these ideas on flow and delivery.

Three perspectives on lean

Lean can be thought of in three related but distinct ways. We will touch on all three perspectives in this chapter:

1. ***Lean is a philosophy of how to run operations.*** It is a coherent set of principles that are founded on smoothing flow through processes by doing all the simple things well, gradually doing them better, meeting customer needs exactly, and squeezing out waste every step of the way.
2. ***Lean is a method of planning and controlling operations.*** Many lean ideas are concerned with how items (materials, information, customers) flow through operations and, more specifically, how operations managers can manage this flow. Uncoordinated flow causes unpredictability, and unpredictability causes waste because people hold inventory, capacity or time to protect themselves against it.
3. ***Lean is a set of tools that improve operations performance.*** The 'engine room' of the lean philosophy is a collection of improvement tools and techniques that are the means for cutting out waste. What is important to understand is how the introduction of lean as a philosophy helped to shift the focus of operations management generally towards viewing improvement as its main purpose.

Case example

Where it came from – Toyota[1]

It would be misleading to neglect describing the Toyota Motor Company's development of lean ideas, because it was seen as the main originator of the lean approach. It was Toyota that developed the set of practices that shaped what we now call 'lean', but which Toyota calls the Toyota Production System (TPS). TPS has two themes, 'just-in-time' and 'jidoka'. Just-in-time is defined as the rapid and coordinated movement of parts throughout the production system and supply network to meet customer demand. It is operationalised by means of *heijunka* (levelling and smoothing the flow of items), *kanban* (signalling to the preceding process that more parts are needed) and *nagare* (laying out processes to achieve smoother flow of parts throughout the production process). Jidoka is described as 'humanising the interface between operator and machine'. Toyota's philosophy is that the machine is there to serve the operator's purpose. The operator should be left free to exercise their judgement. Jidoka is operationalised by means of a fail-safe (or machine jidoka), line-stop authority (or human jidoka) and visual control (at-a-glance status of production processes and visibility of process standards). Toyota believes that both just-in-time and jidoka should be applied rigorously to eliminate waste, where waste is defined as 'anything other than the minimum amount of equipment, items, parts and workers that are absolutely essential to production'.

Cum Okolo/Alamy Stock Photo

Authorities on Toyota also stress the differences between the tools and practices used with Toyota operations and the overall philosophy of its approach to lean synchronisation. What has been called the paradox of the Toyota Production System is that activities, connections and production flows in a Toyota factory are rigidly scripted, yet at the same time Toyota's operations are enormously flexible and adaptable. Activities and processes are constantly being challenged and pushed to a higher level of performance, enabling the company to continually innovate and improve.

It is important to be clear on the distinction between the aim (lean synchronisation), the approach to overcoming the barriers to achieving lean synchronisation, the methods of eliminating waste, and the various techniques that can be used to help eliminate waste. The relationship between these elements is shown in Figure 11.4.

11.2 Diagnostic question: *What are the barriers to lean synchronisation?*

The aim of pure lean synchronisation represents an ideal of smooth, uninterrupted flow without delay, waste or imperfection of any kind. The supply and demand between stages in each process, between processes in each operation and between operations in each supply network are all perfectly synchronised. It represents the ultimate in what customers are looking for from an

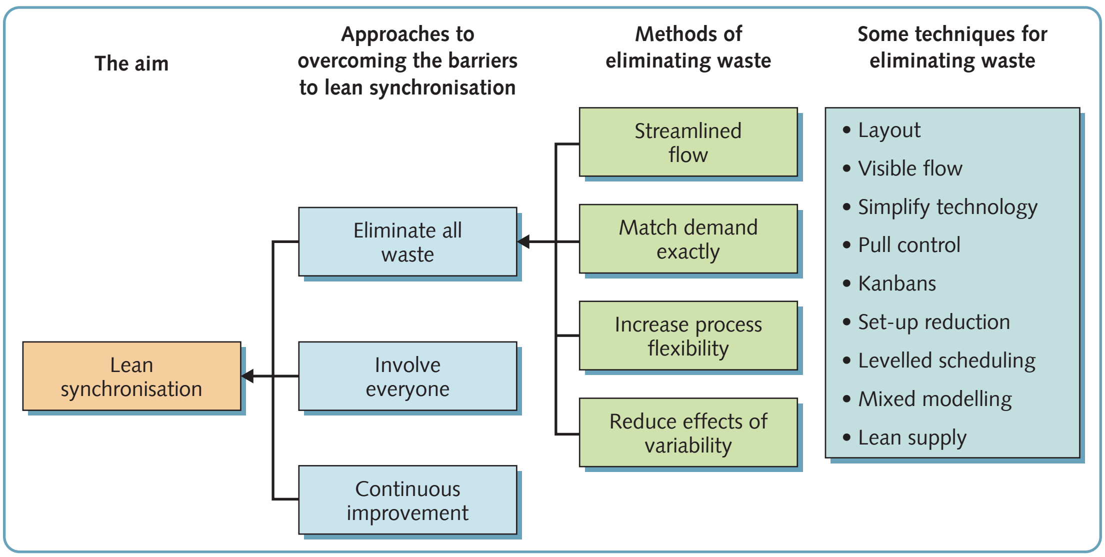


Figure 11.4 **Schematic of the issues covered in this chapter**

operation. But first one must identify the barriers to achieving this ideal state. We group these under three headings:

1. A failure to eliminate waste in all parts of the operation.
2. A failure to harness the contribution of all the people within the operation.
3. A failure to establish improvement as a continuous activity.

The waste elimination barrier

Arguably the most significant part of the lean philosophy is its focus on the elimination of all forms of waste. Waste can be defined as any activity that does not add value. For example, an early classic study by Cummins, the engine company, showed that, at best, an engine was only being worked on for 15 per cent of the time it was in the factory.[2] At worst, this fell to 9 per cent, which meant that for 91 per cent of its time, the operation was adding cost to the engine, not adding value. Although Cummins was already an efficient manufacturer by the standards of the day, the results alerted the company to the enormous waste that still lay dormant in its operations, and which no performance measure then in use had exposed. Cummins shifted its objectives to reducing the wasteful activities and to enriching the value-added ones. Exactly the same phenomenon applies in service processes. Relatively simple requests, such as applying for a driving licence, may only take a few minutes to actually process, yet take days (or weeks) to be returned.

> OPERATIONS PRINCIPLE
> *Focusing on synchronous flow exposes sources of waste.*

Identifying waste is the first step towards eliminating it. There are several categorisations of waste in use. Toyota has described seven types: waste from over-production, waste from waiting time, transportation waste, inventory waste, processing waste, waste of motion and waste from product defects. Here we consolidate these into four broad categories of waste that apply in many different types of operation.

Waste from irregular flow

Perfect synchronisation means smooth and even flow through processes, operations and supply networks. Barriers that prevent streamlined flow include the following:

- *Waiting time* – machine efficiency and labour efficiency are two popular measures that are widely used to measure machine and labour waiting time respectively. Less obvious is the time when products, customers or information wait as inventory or queues, there simply to keep operators busy.
- *Transport* – moving items or customers around the operation, together with double and triple handling, does not add value. Layout changes that bring processes closer together, improvements in transport methods and workplace organisation can all reduce transport waste.
- *Process inefficiencies* – the process itself may be a source of waste. Some operations may only exist because of poor component design, or poor maintenance, and so could be eliminated.
- *Inventory* – regardless of type (product, customer, information), all inventories should become a target for elimination. However, it is only by tackling the causes of inventory or queues, such as irregular flow, that it can be reduced.
- *Wasted motions* – an operator may look busy but sometimes no value is being added by the work. Simplification of work is a rich source of reduction in the waste of motion.

Waste from inexact supply

Perfect synchronisation also means supplying exactly what is wanted, exactly when it is needed. Any under- or over-supply and any early or late delivery will result in waste, something we have already explored in the capacity management chapter (Chapter 8) in particular. Barriers to achieving an exact match between supply and demand include the following:

- *Over-production or under-production* – supplying more, or less, than is immediately needed by the next stage, process or operation.
- *Early or late delivery* – items should only arrive exactly when they are needed. Early delivery is as wasteful as late delivery.
- *Inventory* – again, all inventories should become a target for elimination. However, it is only by tackling the causes of inventory, such as inexact supply, that it can be reduced.

Waste from inflexible response

Customer needs can vary in terms of what they want, how much they want and when they want it. However, processes usually find it more convenient to change what they do relatively infrequently, because every change implies some kind of cost. That is why hospitals schedule specialist clinics only at particular times, and why machines often make a batch of similar products together. Responding to customer demands exactly and instantaneously requires a high degree of process flexibility. Symptoms of inadequate flexibility include the following:

- *Large batches* – sending a large batch of items through a process inevitably increases inventory as the batch moves through the whole process.
- *Delays between activities* – the longer the time (and the cost) of changing over from one activity to another, the more difficult it is to synchronise flow to match customer demand instantaneously.
- *More variation in activity mix than in customer demand* – if the mix of activities in different time periods varies more than customer demand varies, then some 'batching' of activities must be taking place.

Waste from variability

Synchronisation implies exact levels of quality. If there is variability in quality levels then customers will not consider themselves as being adequately supplied. Variability therefore is a significant barrier to achieving synchronised supply. Symptoms of variability include the following:

- *Poor reliability of equipment* – unreliable equipment usually indicates a lack of conformance in quality levels. It also means that there will be irregularity in supplying customers. Either way, it prevents synchronisation of supply.
- *Defective products or services* – waste caused by poor quality is significant in most operations. Service or product errors cause both customers and processes to waste time until they are corrected.

Muda, Mura, Muri

An alternative, but common, classification of waste is the use of the Japanese terms *Muda, Mura and Muri*:

- *Muda* – describes activities in a process that are wasteful because they do not add value to the operation or the customer. The main cause is likely to be poorly communicated objectives (including not understanding the customer's requirements), or the inefficient use of resources. The implication of this is that for an activity to be effective, it must be properly recorded and communicated to whoever is performing it.

- *Mura* – means 'lack of consistency' or unevenness that results in periodic overloading of staff or equipment. For example, if activities are not properly documented so that different people at different times perform a task differently then, unsurprisingly, the result of the activity may be different. The negative effects of this are similar to a lack of dependability.
- *Muri* – means absurd or unreasonable. It is based on the idea that unnecessary or unreasonable requirements put on a process will result in poor outcomes. The implication of this is that appropriate skills, effective planning, accurate estimation of times and schedules will avoid this *muri* overloading waste. In other words, waste can be caused by failing to carry out basic operations planning tasks such as prioritising activities (sequencing), understanding the necessary time (scheduling) and resources (loading) to perform activities. (All these issues are discussed in Chapter 10.)

These three causes of waste are obviously related. When a process is inconsistent (*mura*), it can lead to the overburdening of equipment and people (muri) that, in turn, will cause all kinds of non-value-adding activities (*muda*).

Case example

No more 'faffing around' in the kitchen[3]

(Example written and supplied by Janina Aarts and Mattia Bianchi, Department of Management and Organization, Stockholm School of Economics.)

Cooking can be time-consuming. Knowing this, the celebrity chef Jamie Oliver wrote a book – *Jamie's 30-Minute Meals* – whose philosophy is that cooking a delicious meal should be as quick as and cheaper than buying and heating a takeaway. The book presents 50 ready-made menus with three to four courses per menu, designed to take no more than 30 minutes to prepare. To achieve this, Jamie has, perhaps inadvertently, applied the principles and methods of lean synchronisation to the everyday activity of cooking. Imagine you are cooking a multi-course meal with chicken, rice, salad and a dessert. Traditionally, you would search and look up four different recipes, one for each dish. Because all recipes come from different places, you need to work out the quantity of food to buy, doing the maths in case of shared ingredients across the dishes, how to allocate pots, pans and other equipment to the different ingredients and, most importantly, you need to work out in what order to prepare things, especially if you want all your dishes ready at the same time. Jamie's approach significantly reduces this complexity by ensuring dishes are prepared right when the next step in the process needs it, regardless of which dish it is. In other words, dishes are not cooked in sequence, one after another, but they are prepared and completed simultaneously.

If we identify all the tasks related to preparing the salad (e.g. chopping the vegetables) with the letter A, cooking the rice (e.g. boiling) with letter B, cooking the chicken with letter C, and finally making the dessert with the letter D, then in the traditional way of cooking our task scheduling would look something like AAAA BBBBBBB CCCCCCC DDDD. This results in batching, waiting time and causing dishes to be ready before the dinner is supposed to be served. Conversely, Jamie Oliver's 30-minute cooking involves scheduling tasks in a sequence such as ABCDACBADCBABDC, where single tasks related to different dishes follow smoothly – as the chef chops a salad ingredient, then boils the rice, then chops some more salad ingredients while the chicken is being roasted in the oven and a part of the dessert is being prepared. This way, all dishes are ready at the same time, just in time, and nothing is prepared before it has to be, avoiding any form of waste. Such a levelled approach to scheduling is called *heijunka* (mixed modelling) in lean synchronisation.

In addition, Jamie's lean cooking builds on reduced setup times. At the beginning of each recipe, the equipment needed to prepare the menu is presented under the headline 'To Start'. Other necessary preparation, such as heating the oven, is also specified. Having all equipment ready from the start saves time in the process and is, according to Jamie, a prerequisite for getting everything done in 30 minutes. The use of simple equipment that is suitable for many different purposes also makes the process quicker as changeovers are minimised. The rationale is to make the most out of the time available, eliminating the 'faffing around' in cooking (non-value-added activity in OM language!) and leaving only what is strictly 'good, fast cooking', without compromising on quality.

Short-term pain for a long-term gain

A paradox in the lean synchronisation concept is that adoption may mean some sacrifice of capacity utilisation. In organisations that place a high value on the utilisation of capacity, this can prove particularly difficult to accept. But it is necessary. Return to the process shown in Figure 11.2. When stoppages occur in the traditional system, the buffers allow each stage to continue working and thus achieve high capacity utilisation. But the high utilisation does not necessarily make the system as a whole produce more parts or process more customers. Often the extra processing goes into the large buffer inventories or queues. In a synchronised lean process, any stoppage will affect the rest of the system, causing stoppages throughout the operation. This will necessarily lead to lower capacity utilisation, at least in the short term. However, there is no point in processing products, services or information for the sake of it. Unless the output is useful and enables the operation as a whole to complete all required tasks, there is no point in continuing to work. In fact, processing just to keep utilisation high is not only pointless, it is counterproductive, because the extra inventory or queues that are created merely serve to make improvements less likely. Figure 11.5 illustrates the two approaches to capacity utilisation.

> OPERATIONS PRINCIPLE
> ***Focusing on lean synchronisation can initially reduce resource utilisation.***

The involvement barrier

An organisational culture that supports lean synchronisation must place a very significant emphasis on involving everyone in the organisation. This approach to people management is seen by some as the most controversial aspect of the lean philosophy. It encourages (and often requires) team-based problem solving, job enrichment, job rotation and multi-skilling. The intention is to encourage a high degree of personal responsibility, engagement and 'ownership' of the job. So-called 'basic working practices' are held to be the basic preparation of the operation and its employees for implementing lean synchronisation. They include the following:

- *Discipline* – work standards that are critical for the safety of staff, the environment and quality must be followed by everyone all the time.
- *Flexibility* – it should be possible to expand responsibilities to the extent of people's capabilities. This applies as equally to managers as it does to shop-floor personnel.

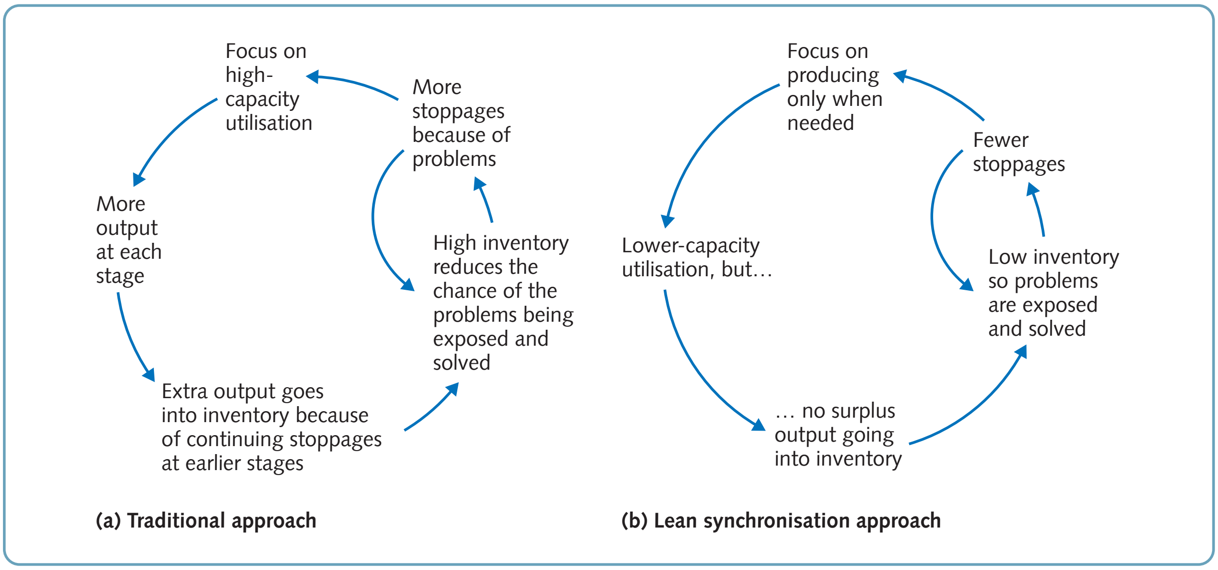


Figure 11.5 The different views of capacity utilisation in (a) traditional and (b) lean synchronisation approaches to planning and controlling flow

- *Equality* – unfair and divisive personnel policies should be discarded. Many companies implement the egalitarian message through to company uniforms, consistent pay structures that do not differentiate between full-time staff and hourly-rated staff, and open-plan offices.
- *Autonomy* – delegate responsibility to people involved in direct activities so that management's task becomes one of supporting processes. Delegation includes giving staff the responsibility for stopping processes in the event of problems, scheduling work, gathering performance monitoring data and general problem solving.
- *Development of personnel* – over time, the aim is to create more company members who can support the rigours of being competitive.
- *Quality of working life (QWL)* – this may include, for example, involvement in decision-making, security of employment, enjoyment and working-area facilities.
- *Creativity* – this is one of the indispensable elements of motivation. Creativity in this context means not just doing a job, but also improving how it is done and building the improvement into the process.
- *Total people involvement* – staff take on more responsibility, such as the selection of new recruits, dealing directly with suppliers and customers over schedules, quality issues, delivery information, spending improvement budgets, and planning and reviewing work done each day.

The concept of continuous learning is also central to the 'involvement of everyone' principle. For example, Toyota's approach to involving its employees includes using a learning method that allows employees to discover the Toyota Production System rules through problem solving.

Case example

Pixar adopts lean[4]

It seems that lean principles (or some lean principles) can be applied even to the most unlikely of processes. None less likely than Pixar Animation Studios, the Academy Award-winning computer animation studio and makers of feature films that have resulted in an unprecedented streak of both critical and box office success, including *Toy Story* (1, 2, 3 and 4), *A Bug's Life*, *Monsters, Inc.*, *Inside Out*, *Finding Nemo*, *The Incredibles*, *Ratatouille*, WALL-E and *Up*. Since its incorporation, Pixar has been responsible for many important breakthroughs in the application of computer graphics (CG) for film-making. So, the company has attracted some of the world's finest technical, creative and production talent in the area. And such 'knowledge-based' talent is notoriously difficult to manage; certainly not the type of process that is generally seen as being appropriate for lean. Managing creativity involves a difficult trade-off between encouraging the freedom to produce novel ideas, yet making sure that they work within an effective overall structure.

Nevertheless, Pixar did get the inspiration from Toyota and the way it uses lean production – in particular, the way Toyota has encouraged continuous advice and criticism from its production line workers to improve its performance. Pixar realised that it could do the same with producing cartoon characters. Adopting constant feedback surfaces problems before they become crises, and provides creative teams with inspiration and challenge. Pixar also devotes a great deal of effort into persuading its creative staff to work together. In similar companies, people may collaborate on specific projects, but are less good at focusing on what's going on elsewhere in the business. Pixar, however, tries to cultivate a sense of collective responsibility. Staff even show unfinished work to one another in daily meetings, thereby becoming used to giving and receiving constructive criticism.

Disney/Pixar/Kobal/Shutterstock

The continuous improvement barrier

Lean synchronisation objectives are often expressed as ideals, such as our previous definition: 'to meet demand instantaneously, with perfect quality and no waste'. While any operation's current performance may be far removed from such ideals, a fundamental lean belief is that it is possible to get closer to them over time. Without such beliefs to drive progress, lean proponents claim improvement is more likely to be transitory than continuous. This is why the concept of continuous improvement is such an important part of the lean philosophy. If its aims are set in terms of ideals that individual organisations may never fully achieve, then the emphasis must be on the way in which an organisation moves closer to the ideal state. The Japanese word that incorporates the idea of continuous improvement is *kaizen*. It is one of the main pillars of process improvement and is explained fully in Chapter 12.

Techniques to address the four sources of waste

Of the three barriers to achieving lean synchronisation (reduce waste, involve everyone and adopt continuous improvement), the last two are addressed further in Chapter 12. The rest of this chapter is therefore devoted to what could be called the 'core of lean synchronisation', which are the various means of cutting out waste. We group these under four main headings: streamlining flow; matching demand exactly; increasing process flexibility; and reducing the effects of variability.

11.3 Diagnostic question: *Is flow streamlined?*

The smooth flow of materials, information and people in the operation is a central idea of lean synchronisation. Long process routes provide opportunities for delay and inventory/queue build-up, add no value and slow down throughput time. So, the first contribution any operation can make to streamline flow is to reconsider the basic layout of its processes. Primarily, reconfiguring the layout of a process to aid lean synchronisation involves moving it down the 'natural diagonal' of process design that was discussed in Chapter 5. Broadly speaking, this means moving from functional layouts towards cell-based layouts, or from cell-based layouts towards line layouts. Either way, it is necessary to move towards a layout that brings more systematisation and control to the process flow. At a more detailed level, typical layout techniques include preventing inventory or queues from building up by restricting space and arranging workstations in such a way that all those who contribute to a common activity are in sight of each other and can provide mutual help. For example, at the Virginia Mason Medical Centre, Seattle, USA, a leading proponent of lean synchronisation in healthcare, many of the waiting rooms have been significantly reduced in their capacity or removed entirely. This forces a focus on the flow of the whole process because patients have literally nowhere to be stored.

> **OPERATIONS PRINCIPLE**
> ***Simple, transparent flow exposes sources of waste.***

Examine the shape of process flow

The pattern that flow makes within or between processes is not a trivial issue. Processes that have adopted the practice of curving line arrangements into U-shaped or 'serpentine' arrangements can have a number of advantages (U-shapes are usually used for shorter lines and serpentines for longer lines). One authority sees the advantages of this type of flow pattern as: staffing flexibility and balance, because the U-shape enables one person to tend to several activities; rework, because it is easy to return faulty work to an earlier station; free flow, because long straight lines interfere with cross travel in the rest of the operation; and teamwork, because the shape encourages a team feeling.

Case example

Kanban visible control at Torchbox web designers[5]

Torchbox is an independently owned web design and development company based in Oxfordshire (see the case example in Chapter 1, 'Torchbox: award-winning web designers' for more details). Despite relying largely on its staff's creative capabilities, this does not mean that lean techniques have no place in its operations. On the contrary, it makes full use of the kanban approach to controlling its work as it progresses through the web design process. '*We know that kanban control originated from car manufacturers like Toyota, but our development teams can also benefit from its basic principles*', said Edward Kay, the Head of Production at Torchbox. '*It is a way of scheduling work based on what needs to be produced and what resources are available to produce it with. At Torchbox we use a large magnetic whiteboard (called the "kanban board") to track completed features through each stage of the design process; from discovery through development, design, demo, deployment and on to the finish of the design (called the "done" stage). Each feature has its own paper slip that physically moves across the board, held in place with a magnet. You can't have more features in progress than the number of magnets you have to hold them in place, so the principle is enforced with a physical limitation. When one feature enters the "Done" column, another one can be pulled through into discovery. There's a pulling process, where completing work allows you to start on something new.*'

At the start of every day, the team has a stand-up meeting at the kanban board, where each member quickly runs through what they did the day before, and what they'll do in the coming day. Each developer has a few tokens that they place on the features they're working on. This helps link up the 'big picture' of how a design's features are developing with the 'little picture' of what each developer is working on each day, and helps teams to make sure that all work being done is being tracked across the board. '*One of the big benefits of using kanban*', said Edward Kay, '*is that because we're visualising the steps a feature goes through to be completed, we're able to see where the bottlenecks are that work gets held up on. If we're finding that a project's features keep getting held up in the design stage, we can bring more designers on to the project to widen the bottleneck. Using kanban with feature-driven development helps us constantly deliver value to our clients. This more measured and controlled approach to handling and controlling incoming work helps ensure that every hour we work produces an hour's worth of value. Ultimately, it's all about delivering great products on time and to budget, and kanban is a great tool to help achieve this.*'

Ensure visibility

Appropriate layout also includes the extent to which all movement is transparent to everyone within the process. High visibility of flow makes it easier to recognise potential improvements to flow. It also promotes quality within the process because the more transparent the operation or process, the easier it is for all staff to share in its management and improvement. Problems are more easily detectable and information becomes simple, fast and visual. Visibility measures include the following:

- clearly indicated process routes using signage;
- performance measures clearly displayed in the workplace;
- coloured lights used to indicate stoppages;
- an area devoted to displaying samples of one's own and competitors' process outputs, together with samples of good and defective output;
- visual control systems (e.g. kanbans, discussed later).

An important technique used to ensure flow visibility is the use of simple but highly visual signals to indicate that a problem has occurred, together with operational authority to stop the process. For example, on an assembly line, if an employee detects some kind of quality problem, he or she could activate a signal that illuminates a light (called an 'andon' light) above the workstation and stops the line. Although this may seem to reduce the efficiency of the line, the

> OPERATIONS PRINCIPLE
> *Visual tools can help everybody to see how work is progressing.*

idea is that this loss of efficiency in the short term is less than the accumulated losses of allowing defects to continue on in the process. Unless problems are tackled immediately, they may never be corrected.

Use small-scale simple process technology

There may also be possibilities to encourage smooth streamlined flow through the use of small-scale technologies; that is, using several small units of process technology, rather than one large unit. In a component manufacturer, for example, small machines have several advantages over large ones. First, they can process different products and services simultaneously. For example, in Figure 11.6 one large machine produces a batch of A, followed by a batch of B, followed by a batch of C. However, if three smaller machines are used, they can each produce A, B or C simultaneously. The system is also more robust. If one large machine breaks down, the whole system ceases to operate. If one of the three smaller machines breaks down, it is still operating at two-thirds effectiveness. Small machines are also easily moved, so that layout flexibility is enhanced. However, investment in capacity may increase in total because parallel facilities are needed, so utilisation may be lower.

Examine all elements of throughput time

Throughput time is often taken as a surrogate measure for waste in a process. The longer that items being processed are held in inventory, moved, checked or subject to anything else that does not add value, the longer they take to progress through the process. So, looking at exactly what happens to items within a process is an excellent method of identifying sources of waste.

Value stream mapping (also known as 'end-to-end' system mapping) is a simple but effective approach to understanding the flow of material and information as a product or service has value added as it progresses through a process, operation or supply chain. It visually maps a product or service's 'production path' from start to finish. In doing so, it records not only the direct activities of creating products and services, but also the 'indirect' information systems that support the direct process. It is called 'value stream' mapping because it focuses on value-adding activities and distinguishes between value-adding and non-value-adding activities. It is similar to process mapping (see Chapter 6) but different in four ways:

1. It uses a broader range of information than most process maps.
2. It is usually at a higher level (5–10 activities) than most process maps.

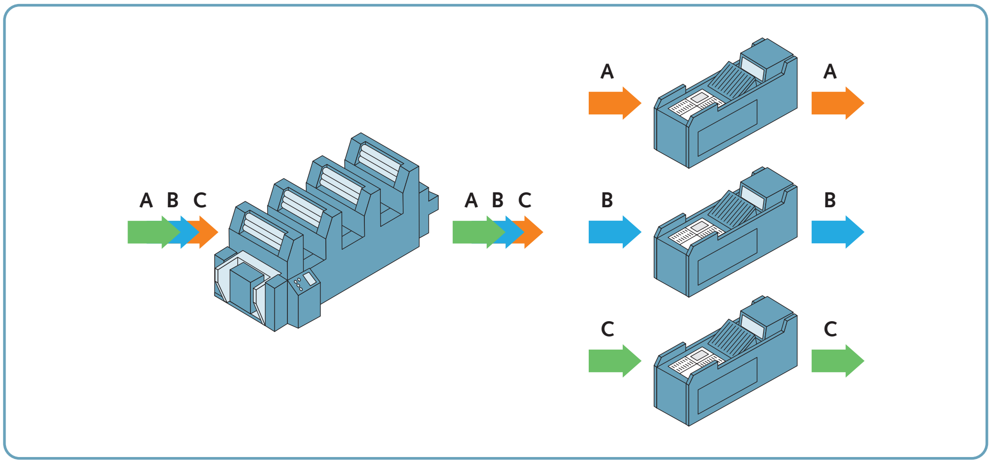


Figure 11.6 **Using several small machines, rather than one large one, allows simultaneous processing, is more robust and is more flexible**

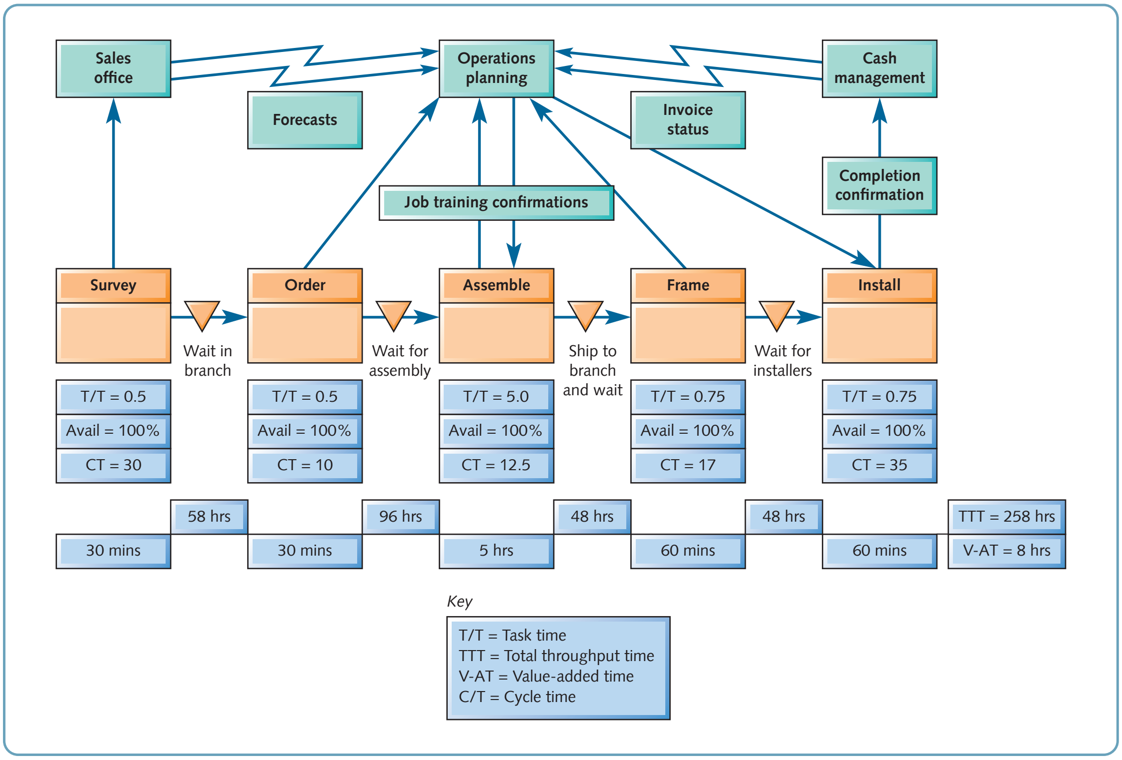


Figure 11.7 Value stream map for an industrial air conditioning installation service

Source: From Slack, N. *Operations Management*, *8e*, © 2016 Pearson Education Limited, UK.

3. It often has a wider scope, frequently spanning the whole supply chain.
4. It can be used to identify where to focus future improvement activities.

A value stream perspective involves working on (and improving) the 'big picture', rather than just optimising individual processes. Value stream mapping is seen by many practitioners as a starting point to help recognise waste and identify its causes. It is a four-step technique that identifies waste and suggests ways in which activities can be streamlined. First, it involves identifying the value stream (the process, operation or supply chain) to map. Second, it involves physically mapping a process, then above it mapping the information flow that enables the process to occur. This is the so-called 'current state' map. Third, problems are diagnosed and changes suggested, making a future state map that represents the improved process, operation or supply chain. Finally, the changes are implemented. Figure 11.7 shows a value stream map for an industrial air conditioning installation service. The service process itself is broken down into five relatively large stages and various items of data for each stage are marked on the chart. The type of data collected here does vary, but all types of value stream maps compare the total throughput time with the amount of value-added time within the larger process. In this case, only eight of the 258 hours of the process are value-adding.

Case example

Andons in Amazon[6]

The principles of an andon cord or light have been applied in Amazon. Every day, service agents at Amazon receive calls from customers who are unhappy with some aspect of the product delivered to them. Customer agents dealing with these complaints are now empowered to make judgements on the extent to which such complaints may be systemic. In cases where they suspect it's a repetitive defect, service agents can 'stop the line' for a particular product. This involves taking the product off the website while the problem is fully investigated. According to Amazon, the improved visibility of the system has eliminated tens of thousands of defects a year and has also given service agents a strong sense of being able to deal effectively with customer complaints. Now an agent can not only refund the individual customer, they can also tell the customer that others won't receive products until the problem has been properly investigated. The firm also claims that around 98 per cent of the times when the andon cord is pulled, there really is a systemic problem, thus highlighting the value of trusting in its service agents to make sensible decisions on when and when not to stop the line.

11.4 Diagnostic question: *Does supply exactly match demand?*

The value of the supply of products or services is always time dependent. Something that is delivered early or late often has less value than something that is delivered exactly when it is needed. We can see many everyday examples of this. For example, parcel delivery companies charge more for guaranteed faster delivery; airlines charge extra for priority boarding; visa applications centres typically offer rapid-turnaround services; and private hospitals charge more by offering faster access to their services. The closer to instantaneous delivery we can get, the more value the delivery has for us and the more we are willing to pay for it. In fact, delivery earlier than it is required can in some cases be just as harmful as late delivery, because it results in information inventories that serve to confuse flow through the process. For example, an Australian tax office used to receive applications by mail, open the mail and send it through to the relevant department which, after processing it, sent it to the next department. This led to piles of unprocessed applications building up within its processes, causing problems in tracing, sorting and prioritising applications, lost applications and, worst of all, long throughput times. Now, mail is opened only when the stages in front are ready and able to process it. Each department requests more work only when it has processed the previous load of work.

OPERATIONS PRINCIPLE
Delivering only and exactly what is needed, when it is needed, smooths flow and exposes waste.

Pull control

The exact matching of supply and demand is often best served by using 'pull control' wherever possible (discussed in Chapter 10). At its simplest, consider how some fast-food restaurants cook and assemble food and place it in the warm area only when the customer-facing server has sold an item. Production is being triggered only by real customer demand. Similarly, supermarkets usually replenish their shelves only when customers have taken sufficient products off the shelf. The movement of goods from the 'back office' store to the shelf is triggered only by the 'empty shelf' demand signal. Some construction companies make it a rule to call for materials deliveries to their sites only the day before those items are actually needed. This not only reduces clutter and the chances of theft, it speeds up throughput time and reduces confusion and inventories. The essence of pull control is to let the downstream stage in a process, operation or supply network pull items through the system rather than have them 'pushed' to them by the supplying stage.

Kanbans

The use of kanbans is one method of operationalising pull control. *Kanban* is the Japanese for card or signal. It is sometimes called the 'invisible conveyor' that controls the transfer of items between the stages of an operation. In its simplest form, it is a card used by a customer stage to instruct its supplier stage to send more items. Kanbans can also take other forms. In some Japanese companies, they are solid plastic markers or even coloured ping-pong balls. Whichever kind of kanban is being used, the principle is always the same; the receipt of a kanban triggers the movement, production or supply of one unit or a standard container of units. If two kanbans are received, this triggers the movement, production or supply of two units or standard containers of units, and so on. Kanbans are the only means by which movement, production or supply can be authorised. Some companies use 'kanban squares'. These are marked spaces on the shop floor or bench that are drawn to fit one or more work pieces or containers. Only the existence of an empty square triggers production at the stage that supplies the square. Likewise, 'kanban whiteboards' are increasingly used to 'pull' activity through the service process (see the Torchbox case example earlier in the chapter).

11.5 Diagnostic question: *Are processes flexible?*

Responding exactly and instantaneously to customer demand implies that operations resources need to be sufficiently flexible to change both what they do and how much they do of it, without incurring high costs or long delays. In fact, flexible processes (often with flexible technologies) can significantly enhance smooth and synchronised flow. For example, new publishing technologies allow professors to assemble printed and e-learning course material customised to the needs of individual courses or even individual students. In this case, flexibility is allowing customised small batches to be delivered 'to order'. In another example, a firm of lawyers used to take ten days to prepare its bills, so customers were not asked to pay until ten days after the work had been done. Now it uses a flexible system that updates each customer's account every day, so bills include all work up to the day before the billing date.

> **OPERATIONS PRINCIPLE**
> ***Changeover flexibility reduces waste and smooths flow.***

Reduce setup times

For many technologies, increasing process flexibility means reducing setup times, defined as the time taken to change over the process from one activity to the next. Compare the time it takes you to change the tyre on your car with the time taken by a Formula 1 team. Setup reduction can be achieved by a variety of methods, such as cutting out time taken to search for tools and equipment, the pre-preparation of tasks that delay changeovers and the constant practice of setup routines. The other common approach to setup time reduction is to convert work that was previously performed while the process was stopped (called internal work) to work that is performed while the process is running (called external work).

In a manufacturing context, making this transition from internal to external wok may include making equipment capable of performing all required tasks so that changeovers become a simple adjustment, and using simple devices to facilitate change of equipment – for example, having equipment on castors to reconfigure processes quickly. In a service context, reducing setup times is also critical. For example, airlines can't make money from aircraft that are sitting idle on the ground – called 'running the aircraft hot' in the industry. For many smaller airlines, the biggest barrier to running hot is that their markets are not large enough to justify passenger flights during both day and night. So, in order to avoid aircraft being idle overnight, they must be used in some other way. That was the motive behind Boeing's 737 'Quick Change' (QC) aircraft. Airlines have the flexibility to use it for passenger flights during the day and, with less than a one hour changeover (setup) time, use it as a cargo plane throughout the night. Boeing engineers designed frames that hold entire rows of seats that can glide smoothly on and off the aircraft, allowing 12 seats to be rolled into place at once. When the planes are used for cargo, the seats are simply rolled out and replaced by special cargo containers designed to fit the curve of the fuselage and prevent damage to the interior. Airlines, such as Aloha Airlines which serves Hawaii, value the aircraft's flexibility as it allows them to provide frequent reliable services in both passenger and cargo markets.

11.6 Diagnostic question: *Is variability minimised?*

One of the biggest causes of variability is variation in the quality of items. This is why a discussion of lean synchronisation should always include an evaluation of how quality conformance is ensured within processes. In particular, the principles of statistical process control (SPC) can be used to understand quality variability. Chapter 13 and its supplement on SPC examine this subject, so in this section we shall focus on other causes of variability. The first of these is variability in the mix of products and services moving through processes, operations or supply networks.

> OPERATIONS PRINCIPLE
> ***Variability in product/service quality, or quantity, or timing, acts against smooth flow and waste elimination.***

Level schedules as much as possible

Levelled scheduling (or *heijunka*) means keeping the mix and volume of flow between stages even over time. For example, instead of producing 500 parts in one batch, which would cover the needs for the next three months, levelled scheduling would require the process to make only one piece per hour regularly. Thus, the principle of levelled scheduling is very straightforward; however, the requirements to put it into practice are quite severe, although the benefits resulting from it can be substantial. The move from conventional to levelled scheduling is illustrated in Figure 11.8. Conventionally, if a mix of products or services was required in a time period (usually a month), a batch size would be calculated for each product and the batches produced in some sequence. Figure 11.8(a) shows three products that are produced in a 20-day time period in a production unit:

Quantity of product/service A required = 3,000

Quantity of product/service B required = 1,000

Quantity of product/service C required = 1,000

Batch size of product/service A = 600

Batch size of product/service B = 200

Batch size of product/service C = 200

Starting at day 1, the unit commences delivering product/service A. During day 3, the batch of 600 As is finished and dispatched to the next stage. The batch of Bs is started but is not finished until day 4. The remainder of day 4 is spent making the batch of Cs and both batches are dispatched at the end of that day. The cycle then repeats itself. The consequence of using large batches is, first, that relatively large amounts of inventory (or queues if this is a service setting and we're batching, for example, different patient groups in an out-patient clinic) accumulate within and between the units, and second, that most days are different from one another in terms of what they are expected to produce (in more complex circumstances, no two days would be the same).

Now suppose that the flexibility of the unit could be increased to the point where the batch sizes for the products or service were reduced to a quarter of their previous levels without loss of capacity (see Figure 11.8(b)):

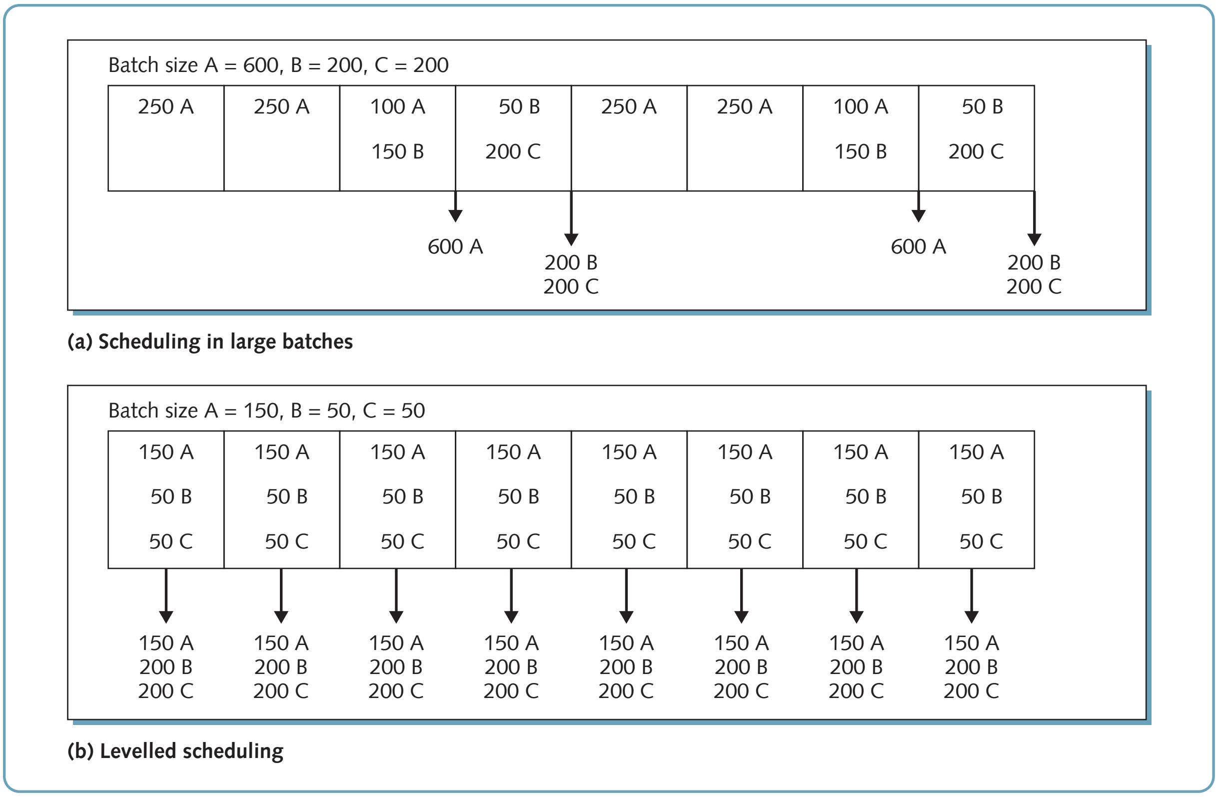


Figure 11.8 **Levelled scheduling equalises the mix of products/services delivered each day**

Batch size of product/service A = 150

Batch size of product/service B = 50

Batch size of product/service C = 50

A batch of each product or service can now be completed in a single day, at the end of which the three batches are dispatched to their next stage. This will reduce the overall level of work-in-progress in the operation. Just as significant, however, is the effect on the regularity and rhythm of the work in each unit. Now every day in the month is the same in terms of what needs to be produced. This makes planning and control of each stage in the operation much easier. For example, if on day 1 of the month the daily batch of As was finished by 11.00 a.m., and all the batches were successfully completed in the day, then the following day the unit will know that, if it again completes all the As by 11.00 a.m., it is on schedule. When every day is different, the simple question 'Are we on schedule to complete our production today?' requires some investigation, but when every day is the same, everyone can tell whether production is on target simply by looking at the clock. Control becomes visible and transparent to all.

Level delivery schedules

A similar concept to levelled scheduling can be applied to many transportation processes. For example, a chain of convenience stores may need to make deliveries of all the different types of products it sells every week. Traditionally, it may have dispatched a truck loaded with one particular product around all its stores so that each store received the appropriate amount of the product that would last it for one week. This is equivalent to the large batches discussed in the previous example. An alternative would be to dispatch smaller quantities of all products in a single truck more frequently. Then, each store would receive smaller deliveries more frequently, inventory levels would be lower and the system could respond to trends in demand more readily because more deliveries means more opportunity to change the quantity delivered to a store. This is illustrated in Figure 11.9.

Adopt mixed modelling where possible

The principle of levelled scheduling can be taken further to give mixed modelling – that is, a repeated mix of outputs. Suppose that the machines in a production unit or employees in a service operation can be made so flexible that they achieve the JIT ideal of a batch size of one. The sequence of individual products or services emerging from the unit could be reduced progressively, as illustrated in Figure 11.10. This would create a steady stream flowing continuously from the unit. However, the sequence does not always fall as conveniently as in Figure 11.10. The working times needed for different products or services are rarely identical and the ratios of required volumes are less convenient. For example, a small business tax return process is required to deal with different tax returns: A (sole traders), B (partnerships) and C (limited companies with 0–3 employees) in the ratio 8:5:4. It could process 800 of A, followed by 500 of B, followed by 400 of A; or 80A, 50B and 40C. But ideally, to sequence the work as smoothly as (theoretically) possible, it would process in the order BACABACABACABACAB . . . repeated. . . repeated . . . etc. Doing this achieves relatively smooth flow (but does rely on significant process flexibility).

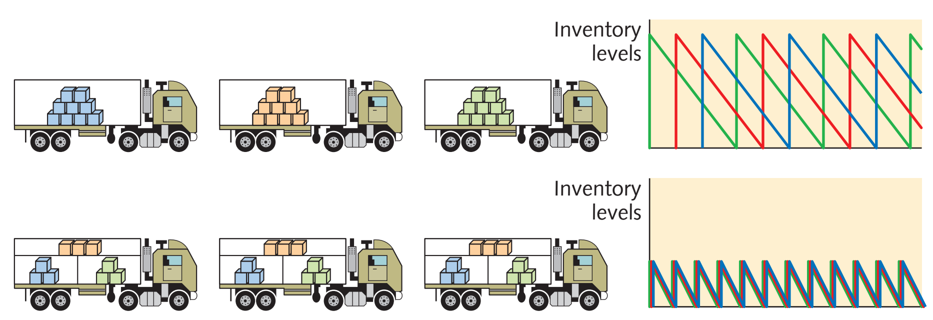
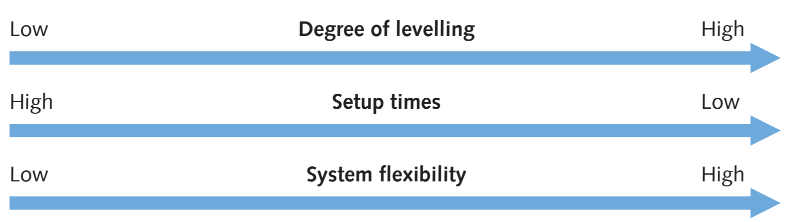


Figure 11.9 Delivering smaller quantities more often can reduce inventory or queue levels



Large batches, e.g.	Small batches, e.g.	Mixed modelling, e.g.
200 A 120 B 80 C	5 A 3 B 2 C	A A B A B C A B C A

Figure 11.10 Levelled scheduling and mixed modelling: mixed modelling becomes possible as the batch size approaches one

Adopt total productive maintenance

Total productive maintenance (TPM) aims to eliminate the variability in operations processes caused by the effect of breakdowns. This is achieved by involving everyone in the search for maintenance improvements. Process owners are encouraged to assume ownership of their equipment and to undertake routine maintenance and simple repair tasks. These principles apply equally to service operations. For example, at a car wash service employees regularly maintain their power-hoses to prevent unnecessary downtime, while university employees may be encouraged to regularly 'clean out' email inboxes, delete old files on their computers and update software with the aim of maintaining system availability speed. By doing so, maintenance specialists can then be freed to develop higher-order skills for improved maintenance systems. (TPM is treated further in Chapter 14.)

11.7 Diagnostic question: *Is lean synchronisation applied throughout the supply network?*

Although most of the concepts and techniques discussed in this chapter are devoted to the management of stages within processes and processes within an operation, the same principles can apply to the whole supply chain. And as any business starts to approach lean synchronisation it will eventually come up against the constraints imposed by the lack of lean synchronisation of the other operations in its supply chain. Ensuring lean synchronisation throughout an entire supply network is clearly a far more demanding task than doing the same within a single process, if only because the nature of the interaction between whole operations is far more complex than between individual stages within a process. A more complex mix of products and services is likely to be involved and the whole network is likely to be subject to a less predictable set of potentially disruptive events. To make a supply chain adopt lean synchronisation means more than making each operation in the chain lean. Rather, one needs to apply the lean synchronisation philosophy to the supply chain as a whole.

> OPERATIONS PRINCIPLE
> ***The advantages of lean synchronisation apply at the level of the process, the operation and the whole supply network.***

Essentially, the principles of lean synchronisation are the same for a supply chain as they are for a process. Fast throughput throughout the whole supply network is still valuable and will save costs throughout the supply network. Lower levels of inventory (product, customer or information) will still make it easier to achieve lean synchronisation. Waste is just as evident (and even larger) at the level of the supply network and reducing waste is therefore a critical task. Streamlined flow, exact matching of supply and demand, enhanced flexibility and minimising variability are all still tasks that will benefit the whole network. The principles of pull control can work between whole operations in the same way as they can between stages within a single process. In fact, the principles and the techniques of lean synchronisation are essentially the same no matter what level of analysis is being used.

Lean supply chains are like an air traffic control system

The concept of the lean supply chain has been likened to an air traffic control system, in that it attempts to provide continuous 'real-time visibility and control' to all elements in the chain. The world's busiest airports handle thousands of departures and arrivals daily. Aircraft are given an identification number and when approaching an airport are detected and contacted. The control tower precisely positions the aircraft in an approach pattern, which it coordinates. The radar detects any small adjustments that are necessary, which are communicated to the aircraft. This real-time visibility and control can optimise airport throughput while maintaining extremely high safety and reliability. Contrast this to how most supply chains are coordinated: information is captured only periodically, and any adjustments to logistics made, output levels adjusted and plans rearranged. Imagine if this was how the airport operated, with only a 'radar snapshot' once a day. Coordinating aircraft with sufficient tolerance to arrange take-offs and landings every two minutes would be out of the question. Aircraft would be jeopardised or, if aircraft were spaced further apart to maintain safety, throughput would be drastically reduced. Yet this is how most supply chains have traditionally operated.

Lean and agile

One continuing debate on how lean principles can be applied across the supply chain concerns whether the supply network should be lean or 'agile'. Professor Martin Christopher of Cranfield University defines 'agility' as 'rapid strategic and operational adaptation to large-scale, unpredictable changes in the business environment. Agility implies responsiveness from one end of the supply chain to the other. It focuses upon eliminating the barriers to quick response, be they organisational or technical.' Other definitions stress that agility is the capability of operating

Case example

Lean construction at St Olav's Hospital, Norway[7]

Construction firms were traditionally slow to adopt lean as a philosophy for managing their operations. Partly, some argued, this was because the characteristics of construction – low volume, high variety – make the principles of lean difficult to implement. However, increasingly, construction firms started embracing lean principles (sometime referred to in construction as 'integrated project delivery'). One example was St Olav's Hospital in Trondheim, central Norway, a regional centre of healthcare expertise. Phase 1 of the hospital's construction had been managed using a traditional project management approach. However, it was not judged a total success, having had several problems including schedule delays and cost overruns. The construction had required substantially more work hours per square metre than had been planned. It was found that too much time had been wasted on non-value-adding work such as transport, waiting, tidying up, etc. Because of these problems, when starting to plan phase 2 a decision was made to actively pursue lean construction practices. After studying how lean had been used in other international projects, phase 2 of the project started, adopting a number of lean principles. These included:

- *Lean engineering and partnering.* For example, the co-location of engineers from different technical disciplines, to encourage the raising of potential problems and ensuring that solutions worked across disciplines, solving problems at the lowest level of the project organisation, co-engineering with contractors using concurrent engineering approaches (see Chapter 3), and applying a 'just-in-time' approach to the production of drawings and engineering documentation.
- *Logistics and purchasing.* For example, keeping the actual construction site as clean and uncluttered as possible, using off-site storage together with a means of transporting materials to the actual site so that the site remained uncluttered, and also using simplified e-purchasing systems that allowed suppliers to deliver materials dependably and fast.
- *Lean construction process.* For example, breaking down the work and planning it with a focus on letting the different disciplines work unimpeded where possible, combating so called 'volume paralysis' (where parts of a project are so large that it is difficult to keep track of progress), structuring planning meetings to ensure that all operations could start when planned, delegating decision-making to those with the best knowledge to solve issues, using quality control throughout the construction process, rather than just at the end, and co-locating the key personnel from all stakeholders.

Because of the adoption of these principles, phase 2 of the St Olav's Hospital project was significantly more successful than phase 1. Despite the increased complexity of phase 2, building time, quality, costs, and health and safety were better than phase 1. There was a 3.4 per cent cost reduction per square metre, improved build quality with 55.1 per cent reduction in warranty costs per square metre, and generally better cooperation between stakeholders.

profitably in a competitive environment of continually changing customer opportunities. The clue lies in how the word 'agile' is often defined; it implies being responsive, quick moving, flexible, nimble, active and constantly ready to change. But some proponents of operational agility go further than this. They see agility as also implying rejection of a planning paradigm that makes any assumption of a predictable future. As with lean, it is more of a philosophy than an approach. Agile encourages a better match to what customers want by placing an emphasis on producing 'emergent' demand as opposed to rigid plans or schedules. Furthermore, rather than uncertainty and change being seen as things to be 'coped with' or preferably avoided, they should be embraced so that agility becomes changing faster than one's customer. Even less-ambitious approaches to agility see it as more than simply organisational flexibility. It involves an organisational mastery of uncertainty and change, where people within the organisation, their capacity to learn from change and their collective knowledge are regarded as the organisation's greatest assets because they allow the operation to respond effectively to uncertainty and change. Continually inventing innovative business process solutions to new market demands becomes a key operations objective.

All this seems very different to the underlying assumptions of the lean philosophy. Again, look at the word: lean means thin, having no superfluous fat. Lean attempts to eliminate waste and provide value to the customer throughout the entire supply chain. It thrives on standardisation, stability, defined processes and repeatability – not at all the way agility has been described. Lean is also a well-defined (although frequently misunderstood) concept. Agility, on the other hand, is a far newer and less 'operationalised' set of relatively strategic objectives, although some operational-level distinctions can be inferred.

The types of principles needed to support a lean philosophy include such things as simple processes, waste elimination, simple (if any) IT and the use of manual and robust planning and control, as well as pull control and kanbans with overall MRP. Agile philosophies, by contrast, require effective demand management to keep close to market needs, a focus on customer relationship management, responsive supply coordination and visibility across the extended supply chain, continuous rescheduling and a quick response to changing demand, short planning cycles, integrated knowledge management and fully exploited e-commerce solutions.

So, are lean and agile philosophies fundamentally opposed? Well, yes and no. Certainly they have differing emphases. Saying that lean equals synchronised, regular flow and low inventory, and agile equals responsiveness, flexibility and fast delivery may be something of a simplification, but it more or less captures the distinction between the two. But because they have different objectives and approaches does not mean that they cannot coexist. Nor does it mean that there is a 'lean versus agile' argument to be resolved. The two approaches may not be complementary, as some consultants claim, but both do belong to the general collection of methodologies that are available to help companies meet the requirements of their markets. In the same way as it was wrong to think that JIT would replace MRP, so agile is not a substitute for lean.[8]

However, agile and lean are each more appropriate for differing market and product/service conditions. Put simply, if product/service variety or complexity is high and demand predictability low, then you have the conditions in which agile principles can keep an operation ready to cope with instability in the business environment. Conversely, if product/service variety is low and demand predictability high, then a lean approach can exploit the stable environment to achieve cost efficiency and dependability. So, the two factors of product/service variety or complexity and demand uncertainty influence whether agile or lean principles should dominate. But what of the conditions where complexity and uncertainty are not related in this manner?

- When complexity is low and demand uncertainty is also low (such as operations that produce commodities), lean planning and control is appropriate.
- When complexity is low and demand uncertainty is high (such as operations that produce fashion-based products/services), agile planning and control is appropriate.

- When complexity is high and demand uncertainty is also high (such as operations that produce 'super value' products/services), project or requirements planning and control (for example MRP II, see Chapter 10) is appropriate.
- When product/service complexity is high and demand uncertainty is low (such as operations that produce 'consumer durable' products/services), a combination of agile and lean planning and control is appropriate.

This last category has been rather clumsily called 'leagile'. Leagile is based on the idea that both lean and agile practices can be employed within supply chains. It envisages an inventory decoupling point that is the separation between the responsive (and therefore agile) 'front end' of the supply chain that reacts fast and flexibly to customer demand, and the efficient (and therefore lean) 'back end' of the supply chain. This is not a new idea and in product-based supply chains involves 'making to forecast' before the decoupling point and 'making (or assembling, adapting or finishing) to order' after it. The idea has many similarities with the idea of 'mass customisation'. In a service context, work that is not for a specific customer, for example a consultancy company creating a 'platform' sector brief, may be worked on in a lean manner, because demand is relatively stable and predictable. By contrast, higher levels of customisation may then be required to adapt such a report to the specific needs of a client, and this may be carried out in an agile manner in the face of changing and often more unpredictable requests from the client.

Critical commentary

- Not surprisingly, various aspects of lean have some critics. After all, it is a concept whose, sometimes counter-intuitive, approach challenges many previously held ideas. Certainly, some operations have had difficulty implementing it successfully, perhaps forgetting that Toyota took decades to develop a fully integrated and coherent philosophy. Lean synchronisation is an aim. It is not something that can be implemented overnight, especially if managers have simply been seduced by its novelty or the bundle of Japanese terms that are part of the concept. Moreover, lean synchronisation principles can be taken to an extreme. When just-in-time ideas first started to have an impact on operations practice in the West, some authorities advocated the reduction of between-process inventories of queues of customers to zero. While, in the long term, this provides the ultimate in motivation to ensure the efficiency and reliability of each process stage, it does not admit the possibility of some processes always being intrinsically less than totally reliable. An alternative view is to allow inventories (albeit small ones) around any process stages with higher-than-average uncertainty. This at least allows some protection for the rest of the system. Even in the best-regulated manufacturing networks, one cannot always account for such events.
- One of the most counter-intuitive issues in lean synchronisation is the way it appears to downplay the idea of capacity underutilisation. And it is true than, when moving towards lean synchronisation, fast throughput time and smooth flow is more important that the high utilisation that can result in build-up of inventory or queues. However, this criticism is not really valid in the long term. Think of the relationship between capacity utilisation and process throughput time (or inventory of products, customers or information), as shown in Figure 11.11. The improvement path envisaged by adopting lean synchronisation is shown as moving from the state that most businesses find themselves in (high utilisation but long throughput times) towards the lean synchronisation ideal (short throughput time). Although, inevitably, this means moving towards a position of lower capacity utilisation, lean synchronisation also stresses a reduction in all types of process variability. As this begins to become reality, the improvement path moves towards the point where throughput time is short and capacity utilisation high. It manages to do this because of the reduction in process variability.

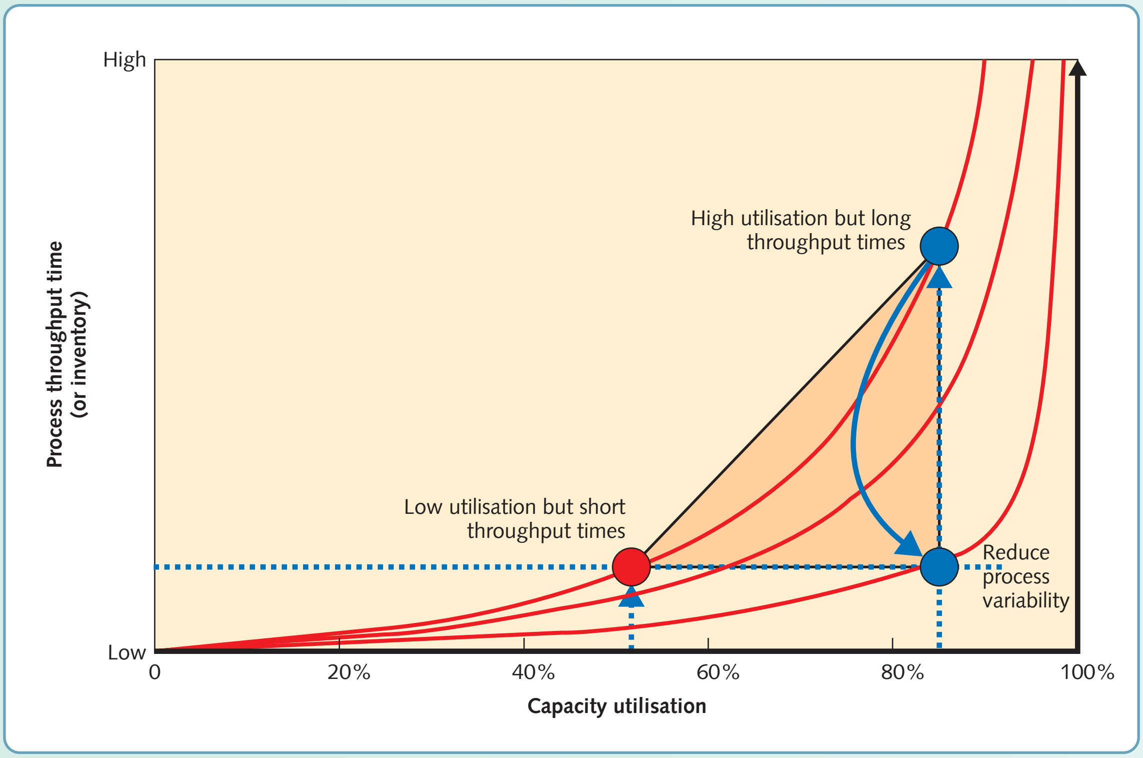


Figure 11.11 Developing lean processes can mean accepting lower utilisation in the short to medium term

- Not all commentators see lean synchronisation-influenced people-management practices as entirely positive. The JIT approach to people management can be viewed as patronising. It may be, to some extent, less autocratic than some Japanese management practices dating from earlier times. However, it is certainly not in line with some of the job design philosophies that place a high emphasis on contribution and commitment. Even in Japan, the JIT approach has not been without its critics. In the early days of its popularity, Kamata wrote an autobiographical description of life as an employee at a Toyota plant called *Japan in the Passing Lane*.[9] His account speaks of 'the inhumanity and the unquestioning adherence' of working under such a system. Some trade union representatives have voiced similar criticisms.
- Any text of this type has to segment the ideas and knowledge contained within its subject in such a way that each set of ideas is explained and communicated as clearly as possible. Yet doing this inevitably means imposing artificial boundaries between the various topics, no more so than in the case of lean synchronisation. There are some particularly evangelical proponents of the lean philosophy who object strongly to separating out the whole concept of lean into a separate chapter. The underlying ideas of lean, they say, have now comprehensively replaced those ideas described as 'traditional' at the beginning of this chapter. Rather, lean principles should be the foundation for the whole of operations and process management. Lean principles have something to tell us about everything in the subject, from quality management to inventory management, from job design to product design. And they are right of course. Nevertheless, the ideas behind lean synchronisation are both counter-intuitive enough and important enough to warrant separate treatment. Also, lean in its pure form is not necessarily equally applicable to every situation (refer to the discussion about lean and agile). Hence the inclusion of this chapter that focuses on this topic. Remember, though, lean synchronisation is one of those topics (like operations strategy, quality and improvement) that has a particularly strong influence over the whole subject of operations and process management.

SUMMARY CHECKLIST

- ☐ Are the benefits of attempting to achieve lean synchronisation well understood within the business?
- ☐ Notwithstanding that the idea derives from manufacturing operations, have the principles been considered for non-manufacturing processes within the business?
- ☐ Is the extent of waste within operations and processes fully understood?
- ☐ Can the flow of items through processes be made more regular?
- ☐ How much inventory (products, customers and information) is building up because of inexact supply?
- ☐ How much waste is caused because of inflexibility in the operation's processes?
- ☐ How much waste is caused because of variability (especially of quality) within the operation's processes?
- ☐ Are capacity utilisation performance measures likely to prove a barrier to achieving lean synchronisation?
- ☐ Does the culture of the organisation encourage the involvement of all people in the organisation in the improvement process?
- ☐ Are the ideas of continuous improvement understood and used in practice?
- ☐ Are the various techniques used to promote lean synchronisation understood and practised?
- ☐ Is the concept of lean synchronisation applied throughout the supply network?
- ☐ Has the possibility of blending push (such as MRP) and pull (such as lean synchronisation) been considered?

Case study

St Bridget's Hospital[10]

When Denize Ahlgren arrived at St Bridget's, one of the main hospitals in the Götenborg area, she knew that it had gained a reputation for fresh thinking on how healthcare could be organised to give better public care at lower cost to the taxpayer. In fact, that was one of the reasons she had taken the job of its Chief of Administration (COA). In particular, Denize had read about St Bridget's 'Quality Care' (QC) initiative. '*Yes, QC is obviously important*', explained Dr Pär Solberg who, in addition to his clinical duties, also headed the QC initiative, '*but don't think that it is only about "quality". We don't just throw money at improving the quality of care; we also want to improve efficiency. Any money saved by improving efficiency can then be invested in improving clinical outcomes*.'

'It all started with quality'

Although run by a private company, St Bridget's is little different from any other Swedish hospital. To its patients, treatment is free following payment of a minimal charge that is universal in Sweden. St Bridget's gets virtually all of its revenue from the government. However, in terms of how it organises itself, it is at the forefront of implementing ideas that are more common in private business. '*It all started with our efforts a few years ago to be systematic in how we measured quality*', said Pär Solberg. '*We felt that quality must be reported on a systematic and logical basis if it is going to be meaningful. It should also be multifaceted, and not just focus on one aspect of quality. We measure three aspects, "reported patient experience" (RPE), what the patient thinks about the total experience of receiving treatment, "reported patient outcome" (RPO), how the patient views the effectiveness of the treatment received, and most importantly "reported clinical outcome" (RCO), how the clinicians view the effectiveness of the treatment. Of course, these three measures are interconnected. So, RPO eventually depends on the medical outcome (RCO) and how much discomfort and pain the treatment triggers. But it is also influenced by the patient's experience (RPE), for example, how well we keep the patient informed, how empathetic our staff are, and so on*.'

'Measuring quality led naturally to continuous improvement'

The hospital's quality measurement processes soon developed into a broader approach to improvement in general. In particular, the idea of continuous improvement began to be discussed. '*Measuring quality led naturally to continuous improvement*', explained Pär Solberg. '*Once we had measurable indicators of quality, we could establish targets, and most importantly we could start to think about what was preventing us from improving quality. This, in turn, led to an understanding of all the processes that affected quality indicators. It was a shift to seeing the hospital as a whole set of processes that governed a set of flows – flows of patients through their treatment stages, flows of clinical staff, flows of information, flows of pharmaceuticals, flows of equipment, and so on. It was a revolution in our thinking. We started examining these flows and looking at how they impacted on our performance and how we could improve the working methods that we considered significant for the quality indicators that we wanted to influence. That was when we discovered the concept of "lean"*.'

'Continuous improvement introduced us to lean'

It was at an 'Improving European Healthcare' conference, which Pär and another colleague attended, that St Bridget's was first introduced to the idea of 'lean'. '*Continuous improvement introduced us to lean. We were talking to some representatives from the UK's National Health Service Institute, who had been involved in introducing lean principles in UK hospitals. They explained that lean was an improvement approach that improved flow and eliminated waste, which had been used successfully in some hospitals to build on continuous improvement. Lean, they said, as developed by Toyota, was about getting the right things to the right place, at the right time, in the right quantities, while minimising waste and being flexible and open to change. It sounded worth following up. However, they admitted that not every attempt to introduce lean principles had met with success*.'

'It can easily all get political'

Intrigued by the conversation, Pär contacted one of the hospitals in the UK that had been mentioned and talked to Marie Watson, who had been the 'Head of Lean' and had initiated several lean projects. She said that one of the problems she had faced was her Chief Executive's insistence on bringing in several firms of consultants to implement lean ideas. To make matters more confusing, when a new Chief Executive was appointed, he brought in his own preferred consultants in addition to those already operating in the hospital. Marie had not been happy with the change. '*Before the change of executives, we had a very clear way of how we were going to move forward and spread lean throughout the organisation, then we became far less clear. The emphasis shifted to get some quick results. But that wasn't why we were set up. Originally it was about having a positive impact, getting people involved in lean, engaging and empowering them towards continuous improvement; there were things that were measurable but then it changed to "show us some quick results". People were forgetting the cultural side of it. Also it can easily all* *get "political". The different consultancy teams and the internal lean initiatives, all had their own territories. For example, we [Marie's internal team] were about to start a study of A&E activities, when they were told to keep away from A&E so as not to "step on the toes" of the firm of consultants working there*.'

'We're not making cars, people are different'

Pär was determined not to make the same mistakes that Marie's hospital had, and consulted widely before attempting any lean improvements with his colleagues. Some were sceptical: '*We're not making cars, people are different and the processes that we put people through repeatedly are more complicated than the processes that you go through to make a car*.' Also, some senior staff were dubious about changes that they perceived would threaten their professional status. Instead of doctors and nurses maintaining separate and defined roles that focused solely on their fields of medical expertise, they were encouraged to work (and sit) together in teams. The teams were also made responsible for suggesting process improvements. But most could be converted. One senior clinician, at first, claimed that, *'this is all a load of rubbish. There's no point in mapping this process, we all know what happens: the patient goes from there to there and this is the solution and that's what we need to do*.' Yet only a few days later, he said, '*I never realised this is what really happens, that won't work now will it, actually this has been great because I never understood, I only saw my bit of it, now I understand all of the process*.'

'It works, it makes things better for the patients'

Over time, most (although not quite all) scepticism was overcome, mainly because, in the words of one doctor, '*It works, it makes things better for the patients*'. As more parts of the hospital became convinced of the effectiveness of the lean approach, the improvements to patient flow and quality started to accumulate. Some of the first improvements were relatively simple, such as a change of signage (to stop patients getting lost). Another simply involved a roll of yellow tape. Rather than staff wasting precious time looking for equipment such as defibrillators, the yellow tape was used to mark a spot on the floor where the machines were always kept. Another improvement involved using magnetic dots on a progress chart to follow each patient's progress and indicate which beds were free. Some were even simpler, for example discharging patients throughout the day rather than all at the same time, so that they could easily find a taxi. Other improvements involved more analysis, such as reducing the levels of stock being held (e.g. 25,000 pairs of surgical gloves from 500 different suppliers). Some involved a complete change in assumptions, such as the effectiveness of the medical records department: '*It was amazing. We just exploded the myth that when you didn't get case notes in a clinical area it was medical records' fault. But it never was. Medics had notes in their cars, they had them at home, we had a thousand notes in the secretaries' offices, there were notes in wards, in drawers and cupboards, they were all over the place. And we wondered why we couldn't get case notes! Two people walked seven miles a day just to find notes for cases!*' (Pär Solberg)

'We need to go to the next level'

Denize Ahlgren was understandably impressed by the improvements that Pär had outlined to her; however, Pär was surprisingly downbeat about the future. '*OK, I admit that we have had some impressive gains from continuous improvement and latterly from the adoption of lean principles. I am especially impressed with Toyota's concept of the seven types of waste*' [for more details see the case example 'Where it came from – Toyota' earlier in this chapter]. '*It is both a conceptually powerful and a very practical idea for identifying where we could improve. Also, the staff like it. But it's all getting like a box-ticking exercise. Looking for waste is not exactly an exciting or radical idea. The more that I study how lean got going in Toyota and other manufacturing plants, the more I see that we haven't really embraced the whole philosophy. Yet, at the same time, I'm not totally convinced that we can. Perhaps some of the doubters were right; a hospital isn't a car plant, and we can apply only some lean ideas*.'

Ironically, as Pär was having doubts, some of his colleagues were straining to do more. One clinician, in particular, Fredrik Olsen, Chief Physician at St Bridget's lower-back pain clinic, thought that his clinic could benefit from a more radical approach: '*We need to go to the next level. The whole of Toyota's philosophy is concerned with smooth synchronous flow, yet we haven't fully got our heads round that here. I know that we are reluctant to talk about "inventories" of patients, but that is exactly what waiting rooms are. They are "stocks" of people, and we use them in exactly the same way as pre-lean manufacturers did – to buffer against short-term mismatches between supply and demand. What we should be doing is tackling the root causes of the mismatch. Waiting rooms are stopping us from moving towards smooth, value-added flow for our patients*.'

Fredrik went on to make what Denize thought was an interesting, but radical, proposal. He proposed scrapping

Photo_Concepts/Cultura/Getty Images

the current waiting room for the lower-back pain clinic and replacing it with two extra consulting rooms to add to the two existing consulting rooms. Patients would be given appointments for specific times rather than being asked to arrive 'on the hour' (effectively in batches), as at present. A nurse would take the patients' details and perform some preliminary tests, after which they would call in the specialist physician. Staffing levels during clinic times would be controlled by a nurse who would also monitor patient arrival, direct them to consulting rooms and arrange any follow-up appointments (for MRI scans, for example).

Denize was not sure about Fredrik's proposal. '*It seems as though it might be a step too far. Patients expect to wait until a doctor can see them, so I'm not sure what benefits would result from the proposal. And what is the point of equipping two new consulting rooms if they are not going to be fully utilised?*'

Questions

1. What benefits did St Bridget's get from adopting first a continuous improvement and then a lean approach?
2. Do you think that Pär Solberg is right in thinking that there is a limit to how far a hospital can go in adopting lean ideas?
3. On the St Bridget's website there are several references to its 'Quality Care' programme, but none to its lean initiatives, even though lean is regarded as important by most clinicians and administrators in the hospital. Why do you think this might be?
4. Denize cannot see the benefits of Fredrik's proposal. What do you think they might be?
5. Are any benefits from scrapping the waiting room in the clinic worth the underutilisation of the four consulting rooms that Fredrik envisages creating?

Applying the principles

Some of these exercises can be answered by reading the chapter. Others will require some general knowledge of business activity and some might require an element of investigation. **Model answers for the first two questions are available on the student companion website**. *For model answers to the other questions in this section, please ask your tutor/lecturer.*

1. Re-examine the description of the Toyota Production System at the beginning of the chapter.

a) List all the different techniques and practices that Toyota adopts. Which of these would you call lean philosophies and which are lean techniques?

b) How are operations objectives (quality, speed, dependability, flexibility, cost) influenced by the practices that Toyota adopts?

2. Consider this record of an ordinary flight: '*Breakfast was a little rushed but left the house at 6.15. Had to return a few minutes later, forgot my passport. Managed to find it and leave (again) by 6.30. Arrived at the airport 7.00, dropped Angela off with bags at terminal and went to the long-stay car park. Eventually found a parking space after 10 minutes. Waited 8 minutes for the courtesy bus. Six-minute journey back to the terminal; we started queuing at the check-in counters by 7.24. Twenty-minute wait. Eventually get to check-in and find that we have been allocated seats at different ends of the plane. Staff helpful but take 8 minutes to sort it out. Wait in queue for security checks for 10 minutes. Security decide I look suspicious and search bags for 3 minutes. Waiting in lounge by 8.05. Spend 1 hour and 5 minutes in lounge reading computer magazine and looking at small plastic souvenirs. Hurrah, flight is called at 9.10, takes 2 minutes to rush to the gate and queue for further 5 minutes at gate. Through the gate and on to air bridge that is a continuous queue going onto plane; takes 4 minutes but finally in seats by 9.21. Wait for plane to fill up with other passengers for 14 minutes. Plane starts to taxi to runway at 9.35. Plane queues for 10 minutes to take off. Plane takes off at 9.45. Smooth flight to Amsterdam, 55 minutes. Stacked in queue of planes waiting to land for 10 minutes. Touch down at Schiphol Airport 10.50. Taxi to terminal and wait 15 minutes to disembark. Disembark at 11.05 and walk to luggage collection (calling at lavatory on way); arrive luggage collection 11.15. Wait for luggage 8 minutes. Through customs (not searched by Netherlands security who decide I look trustworthy) and to taxi rank by 11.26. Wait for taxi for 4 minutes. Into taxi by 11.30; 30-minute ride into Amsterdam. Arrive at hotel 12.00*.'

a) Analyse the journey in terms of value-added time (actually going somewhere) and non-value-added time (the time spent queuing, etc.).

b) Visit the websites of two or three airlines and examine their business-class and first-class services to look for ideas that reduce the non-value-added time for customers who are willing to pay the premium.

c) Next time you go on a journey, time each part of the journey and perform a similar analysis.

3. An insurance underwriting process consists of the following separate stages:

Stage	*Processing time per application (minutes)*	*Average work-in-progress before the stage*
Data entry	30	250
Retrieve client details	5	1,500
Risk assessment	18	300
Inspection	15	150
Policy assessment	20	100
Dispatch proposal	10	100

What is the 'value-added' percentage for the process? (*Hint*: use Little's Law to work out how long applications have to wait at each stage before they are processed. Little's Law is covered in Chapter 6.)

4. Examine the value-added versus non-value-added times for some other services. For example:

a) Handing-in an assignment for marking if you are currently studying for a qualification (the typical elapsed time between handing the assignment in and receiving it back with comments). How much of this elapsed time do you think is value-added time?

b) Posting a letter (the elapsed time between posting the letter in the box and it being delivered to the recipient).

5. Go back to Chapter 10 and re-read the description of the theory of constraints (TOC). Now consider the similarities and differences between TOC and lean synchronisation in terms of their overall objectives, measures of effectiveness, improvement methods and implementation.

Notes on chapter

1 The information on which this example is based is taken from: Spears, S. and Bowen, H.K. (1999) 'Decoding the DNA of the Toyota Production System', *Harvard Business Review*, September–October, pp. 96–106.

2 Lee, D.C. (1987) 'Set-up time reduction: Making JIT work' in C.A. Voss (ed), *Just-in-Time Manufacture*, IFS/Springer-Verlag.

3 Example written and supplied by Janina Aarts and Mattia Bianchi, Department of Management and Organization, Stockholm School of Economics.

4 The information on which this example is based is taken from: Catmull, E. (2008) 'How Pixar fosters collective creativity', *Harvard Business Review*, September, p. 4.

5 The information on which this example is based is taken from: an interview with Edward Kay, Tom Dyson and Olly Willans of Torchbox; Torchbox website, http://www.torchbox.com/ [accessed 27 September 2020]; we are grateful to everyone at Torchbox for their help and for allowing us access to their operation.

6 The information on which this example is based is taken from: Onetto, M. (2014) 'When Toyota met e-commerce: Lean at Amazon', *McKinsey Quarterly*, 1 February; Liker, J. (2004) *The Toyota Way: 14 management principles from the world's greatest manufacturer*, McGraw-Hill Education; Rosenthal, R. (2002) 'The Essence of Jidoka', SME Lean Directions Newsletter.

7 The information on which this example is based is taken from: Banawi, A. and Bilec, M.M. (2014) 'A framework to improve construction processes: Integrating lean, green and Six Sigma', *International Journal of Construction Management*, 14 (1), pp. 45–55; Andersen, B., Belay, A.M. and Amdahl Seim, E. (2012) 'Lean construction practices and its effects: A case study at St Olav's Integrated Hospital, Norway', *Lean Construction Journal*, pp. 122–149, www.leanconstruction.org.uk/lean-construction-journal [accessed 27 September 2020].

8 Kruse, G. (2002) 'IT-enabled lean agility', *Control*, November.

9 Kamata, S. (1983) *Japan in the Passing Lane: An insider's account of life in a Japanese auto factory*, Allen and Unwin.

10 This case is based on the work of several real hospitals, in Scandinavia and the rest of the world, that have used the concepts of lean operations to improve their performance. However, all names and places are fictional and no connection to any specific hospital is intended.

Taking it further

Bicheno, J. and Holweg, M. ***(2016) The Lean Toolbox: The essential guide to lean transformation, 5th edition, Picsie Books.*** *A practical guide from two of the European authorities on all matters lean.*

Holt, P. ***(2019) The Simplicity of Lean: Defeating complexity, delivering excellence, Global Publisher Services.*** *Good practical advice from a distinguished practitioner.*

Holweg, M. ***(2007) 'The genealogy of lean production', Journal of Operations Management, 25 (2), pp. 420–437.*** *An excellent overview of how lean ideas developed.*

Mann, D. ***(2010) Creating a Lean Culture: Tools to sustain lean conversion, 2nd edition, Productivity Press.*** *Treats the soft side of lean.*

Modig, N. and Ahlstrom, P. ***(2012) This is Lean: Resolving the efficiency paradox, Rheologica Publishing.*** *This book provides a very practical guide to what lean is and its application in a variety of sectors. Not only does this book demonstrate a clear understanding of how the various aspects of lean come together, it does this in a very readable way.*

Netland, T.H. and Powell, D.J. ***(eds) (2019) The Routledge Companion to Lean Management, Routledge.*** *An academic but wide-ranging collection of articles on lean.*

Womack, J.P. and Jones, D.T. ***(2003) Lean Thinking: Banish waste and create wealth in your corporation, Free Press.*** *Some of the lessons from The Machine that Changed the World but applied in a broader context.*

12 Improvement

Introduction

Improvement means to make something better. And all operations, no matter how well managed, are capable of being made better. At one time, operations managers were expected simply to 'run the operation', 'keep the show on the road' and 'maintain current performance'. No longer. In fact, in recent years the emphasis has shifted markedly towards improvement as one of the main responsibilities of operations managers. Of course, the whole of this text can be taken as focusing on improving the performance of individual processes, operations and whole supply networks. Yet there are some issues that relate specifically to the activity of improvement itself. This chapter, and the following two (Chapter 13, which deals with quality management and Chapter 14, which deals with risk and resilience) all have a common objective – to enhance operations and process performance. Figure 12.1 shows the position of the ideas described in this chapter within the general model of operations management.

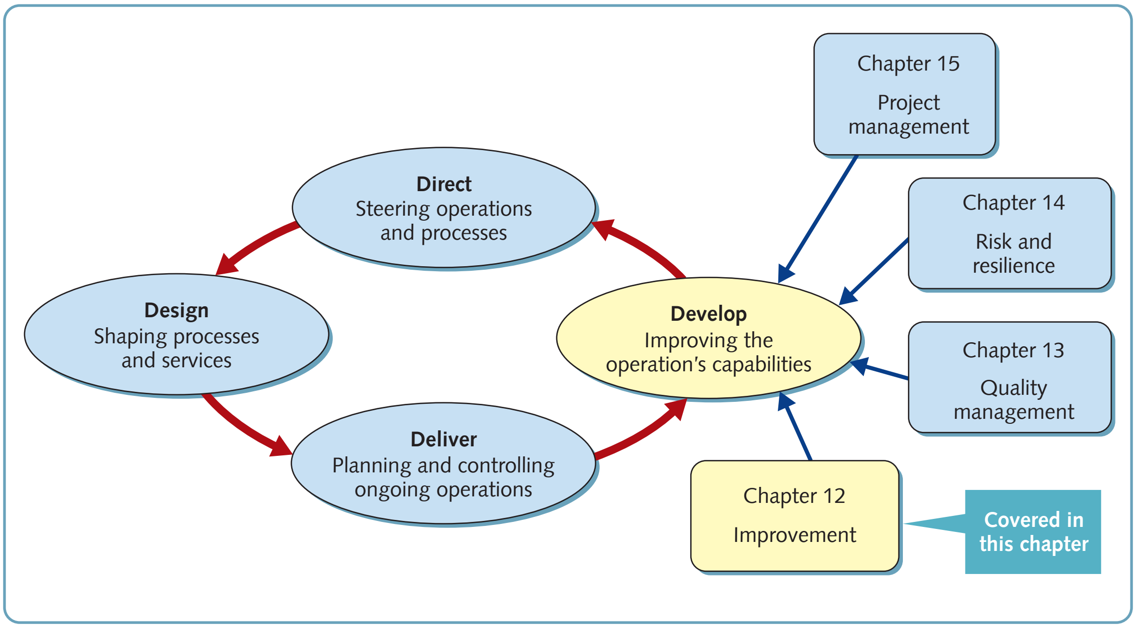


Figure 12.1 Improvement is the activity of closing the gap between the current and the desired performance of an operation or process

EXECUTIVE SUMMARY

12.1 Why is improvement so important?

12.2 What is the gap between current and required performance?

12.3 What is the most appropriate improvement path?

12.4 What techniques could be used to facilitate improvement?

12.5 How can improvement be made to stick?

12.1 Why is improvement so important?

Improvement is the activity of closing the gap between the current and the desired performance of an operation or process. It is increasingly seen as the ultimate objective for all operations and process management activity. Furthermore, almost all popular operations initiatives in recent years, such as total quality management, lean operations, business process re-engineering and Six Sigma, have focused on performance improvement. It involves assessing the gaps between current and required performance, balancing the use of continuous improvement and breakthrough improvement, adopting appropriate improvement techniques and attempting to ensure that the momentum of improvement does not fade over time.

12.2 What is the gap between current and required performance?

Assessing the gap between actual and desired performance is the starting point for most improvement. This requires two sets of activities: first, assessing the operation's and each process's current performance; and second, deciding on an appropriate level of target performance. The first activity will depend on how performance is measured within the operation. This involves deciding what aspects of performance to measure, which are the most important aspects of performance, and what detailed measures should be used to assess each factor. The balanced scorecard is an approach to performance measurement that is currently influential in many organisations. Setting targets for performance can be done in a number of ways. These include historically based targets, strategic targets that reflect strategic objectives, external performance targets that relate to external and/or competitor operations, and absolute performance targets based on the theoretical upper limit of performance. Benchmarking is an important input to this part of performance improvement.

12.3 What is the most appropriate improvement path?

Two improvement paths represent different philosophies of improvement, although both may be appropriate at different times. They are breakthrough improvement and continuous improvement. Breakthrough improvement focuses on major and dramatic changes that are intended to result in dramatic increases in performance. The business process re-engineering approach is typical of breakthrough improvement. Continuous improvement focuses on small but never-ending improvements that become part of normal operations life. Its objective is to make improvement part of the culture of the organisation. Often continuous improvement involves the use of multi-stage improvement cycles for regular problem solving. The Six Sigma approach to improvement brings many existing ideas together and can be seen as a combination of continuous and breakthrough improvement.

12.4 What techniques could be used to facilitate improvement?

Almost all techniques in operations management contribute directly or indirectly to the performance improvement. However, some more general techniques have become popularly associated with improvement. These include scatter diagrams (correlation), cause–effect diagrams, Pareto analysis and why–why analysis.

12.5 How can improvement be made to stick?

One of the biggest problems in improvement is to preserve improvement momentum over time. One factor that inhibits improvement becoming accepted as a regular part of operations activity is the emphasis on the fashionability of each new improvement approach. Most new improvement ideas contain some worthwhile elements but none will provide the ultimate answer. There must be some overall management of the improvement process that can absorb the best of each new idea. And, although authorities differ to some extent, most emphasise the importance of an improvement strategy, top-management support and training.

12.1 Diagnostic question: *Why is improvement so important?*

Most of us would like to get better at the things we value. And businesses are (or should be) no different – they generally want to get better. Not just for the sake of their own excellence, although that may be one factor, but mainly because improving operations performance has such an impact on what any organisation is there to do. Emergency services want to reach distressed people faster and treat them better because by doing so they are fulfilling their role more effectively. Package delivery businesses want to deliver more reliably, at lower cost and reducing emissions because it means happier customers, higher profits and less pollution. Development charities want to target their aid and campaign for improvement in human conditions as wisely and efficiently as possible because more money will find its way to beneficiaries, rather than be wasted or consumed in administration. It is not surprising then that the whole emphasis of operations management has shifted towards emphasising improvement. Operations managers are judged not only on how they meet their ongoing responsibilities of producing products and services to acceptable levels of quality, speed, dependability, flexibility and cost, but also on how they improve the performance of the operations function overall.

Improvement comes from closing the gap between what you are and what you want to be. Or specifically, in an operations context, it comes from closing the gap between current and desired performance. Performance improvement is the ultimate objective of operations and process management. It has also become the subject of innumerable ideas that have been put forward as particularly effective methods of ensuring improvement. These include many that are described in this text – for example, total quality management (TQM), 'lean' operations, business process re-engineering (BPR), Six Sigma, and so on. All of these, and other, ideas have something to contribute. What is important is that all managers develop an understanding of the underlying elements of improvement.

> OPERATIONS PRINCIPLE
> ***Performance improvement is the ultimate objective of operations and process management.***

Why the focus on improvement?

Various reasons have been suggested to explain the shift towards a focus on improvement in professional operations managers' activities:

- There is a perceived increase in the intensity of competitive pressures (or 'value for money' in not-for-profit or public-sector operations). In fact, economists argue about whether markets are really getting more competitive. As far as improvement is concerned it doesn't matter; there is a *perception* of increased competitive pressure, and certainly the owners of operations (shareholders or governments) are less likely to tolerate poor returns or value for money.
- The nature of world trade is changing. Some developing economies are emerging as both producers and consumers of products and services. Politically inspired trade conflicts, alliances and separations have created operating uncertainty. The coronavirus pandemic disrupted previously well-established supply chains. All of which has increased the value of operational flexibility as well as introducing cost pressures in countries with relatively expensive labour and infrastructure costs.
- New technology has introduced opportunities to both improve operations practice and disrupt existing markets. Look at how operations in the music business have had to adapt their working practices to downloading and music streaming.
- The interest in operations improvement has resulted in the development of many new ideas and approaches to improving operations that have, in turn, focused attention on improvement. The more ways there are to improve operations, the more operations will be improved.

The Red Queen effect

The scientist Leigh Van Valen was looking to describe a discovery that he had made while studying marine fossils. He had established that, no matter how long a family of animals had already existed, the probability that the family will become extinct is unaffected. In other words, the struggle for survival never gets easier. However well a species fits with its environment, it can never relax. The analogy that Van Valen drew came from *Through the Looking Glass*, by Lewis Carroll.[1] In the book, Alice had encountered living chess pieces and, in particular, the Red Queen. *'"Well, in our country", said Alice, still panting a little, "you'd generally get to somewhere else – if you ran very fast for a long time, as we've been doing". "A slow sort of country" said the Queen. "Now, here, you see, it takes all the running you can do, to keep in the same place. If you want to get somewhere else, you must run at least twice as fast as that!"'*

In many respects this is like business. Improvements and innovations may be imitated or countered by competitors. For example, in many competitive sectors, such as consumer electronics or the automotive sector, the quality of most firms' products is very significantly better than it was two decades ago. This reflects the improvement in those firms' operations processes. Yet their relative competitive position has in many cases not changed. Those firms that have improved their competitive position have improved their operations performance more than competitors. Where improvement has simply matched competitors, survival has been the main benefit. The implication for operations improvement is clear: it is even more important, especially when competitors are actively improving their operations.

12.2 Diagnostic question: *What is the gap between current and required performance?*

The gap between how an operation or process is currently performing, and how it wishes to perform, is the key driver of any improvement initiative. The wider the gap, the more importance is likely to be given to improvement. But, in order to harness the gap as a driver of improvement, it must be addressed in some detail, both in terms of exactly what is failing to meet targets, and by how much. Answering these questions depends on the operation's ability to do three things: assess its current performance; derive a set of target levels of performance that the organisation can subscribe to; and compare current against target performance in a systematic and graphic manner that demonstrates to everyone the need for improvement.

Assessing current performance – performance measurement

Some kind of *performance measurement* is a prerequisite for judging whether an operation is good, bad or indifferent, although this is not the only reason for investing in effective performance measurement. Without it, it would be impossible to exert any control over an operation on an ongoing basis. However, a performance measurement system that gives no help to ongoing improvement is only partially effective. Performance measurement, as we are treating it here, concerns three generic issues:

- Which factors to include as performance measures.
- Which are the most important performance measures.
- Which detailed measures to use.

> **OPERATIONS PRINCIPLE**
> ***Performance measurement is a prerequisite for the assessment of operations performance.***

Which factors to include as performance measures

An obvious starting point for deciding which performance measures to adopt is to use the five generic performance objectives: quality, speed, dependability, flexibility and cost. These can be broken down into more detailed measures, or they can be aggregated into 'composite' measures, such as 'customer satisfaction', 'overall service level', or 'operations agility'. These composite measures may be further aggregated by using measures such as 'achieve market objectives', 'achieve financial objectives', 'achieve operations objectives' or even 'achieve overall strategic objectives'. The more aggregated performance measures have greater strategic relevance, in so much as they help to draw a picture of the overall performance of the business, although by doing so they necessarily include many influences outside of those that operations performance improvement would normally address. The more detailed performance measures are usually monitored more closely and more often, and although they provide a limited view of an operation's performance, they do provide a more descriptive and complete picture of what should be and what is happening within the operation. In practice, most organisations will choose to use performance targets from throughout the range. This idea is illustrated in Figure 12.2.

Choosing the most important performance measures

One of the problems of devising a useful performance measurement system is trying to achieve some balance between having a few key measures on one hand (straightforward and simple, but may not reflect the full range of organisational objectives), or having many detailed measures (complex and difficult to manage, but capable of conveying many nuances of performance) on the other. Broadly, a compromise is reached by making sure that there is a clear link between the operation's overall strategy, the most important (or 'key') performance indicators (KPIs) that reflect strategic objectives, and the bundle of detailed measures that are used to 'flesh out' each key performance indicator. Obviously, unless strategy is well defined then it is difficult to 'target' a narrow range of key performance indicators.

> OPERATIONS PRINCIPLE
> ***Without strategic clarity, key performance indicators cannot be appropriately targeted.***

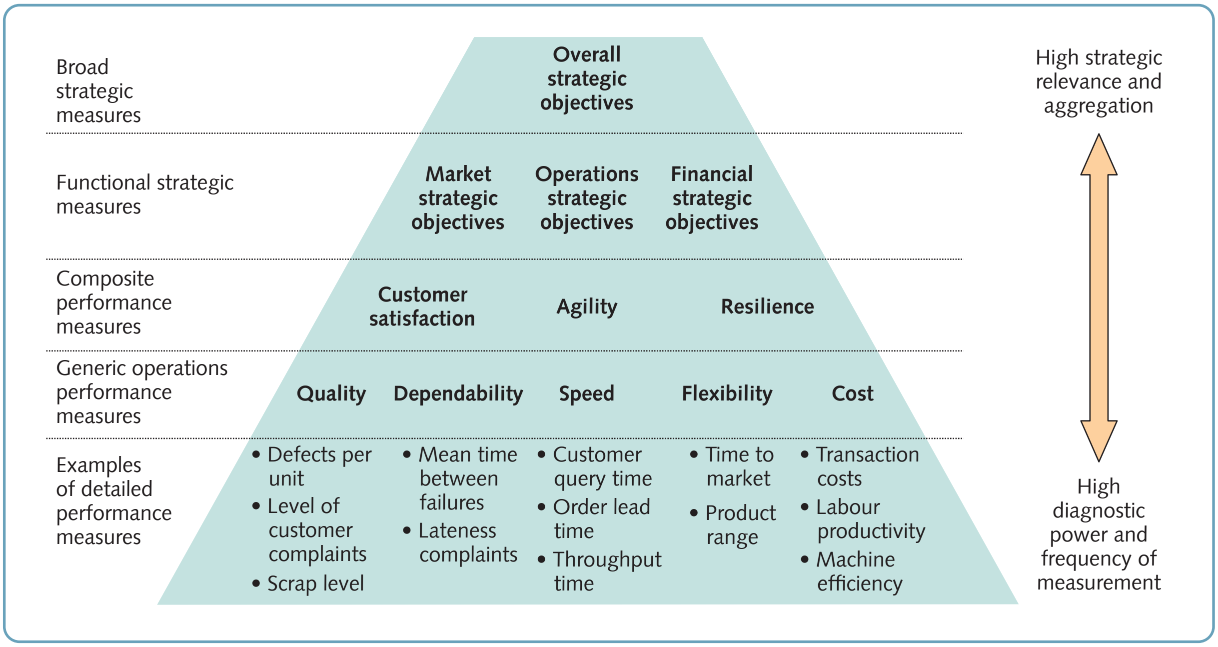


Figure 12.2 **Performance measures can involve different levels of aggregation**

Which detailed measures to use

The five performance objectives – quality, speed, dependability, flexibility and cost – are really composites of many smaller measures. For example, an operation's cost is derived from many factors, which could include the purchasing efficiency of the operation, the efficiency with which it converts materials, the productivity of its staff, the ratio of direct to indirect staff, and so on. All of these measures individually give a partial view of the operation's cost performance, and many of them overlap in terms of the information they include. However, each of them does give a perspective on the cost performance of an operation that could be useful either to identify areas for improvement or to monitor the extent of improvement. If an organisation regards its 'cost' performance as unsatisfactory, disaggregating it into 'purchasing efficiency', 'operations efficiency', 'staff productivity', etc. might help to explain the root cause of the poor performance. Table 12.1 shows some of the partial measures that can be used to judge an operation's performance.

Performance measurement and performance management

A distinction that can cause some confusion is that between performance *measurement* and performance *management*. Performance measurement is the attempt to assess how well (or

Table 12.1 Some typical partial measures of performance

Performance objective	*Some typical measures*
Quality	Number of defects per unit Level of customer complaints Scrap level Warranty claims Mean time between failures Customer satisfaction score
Speed	Customer query time Order lead time Frequency of delivery Actual versus theoretical throughput time Cycle time
Dependability	Percentage of orders delivered late Average lateness of orders Proportion of products in stock Mean deviation from promised arrival Schedule adherence
Flexibility	Time needed to develop new products/services Range of products/services Machine changeover time Average batch size Time to increase activity rate Average capacity/maximum capacity Time to change schedules
Cost	Minimum delivery time/average delivery time Variance against budget Utilisation of resources Labour productivity Added value Efficiency Cost per operation hour

not) an operation is performing. Performance management is all the actions taken to affect performance in some way. So performance management is what operations managers do to achieve the measures of performance that they are aiming for. And if this seems like a subtle distinction, it is. In fact, measurement and management are not separable.[2] The two things follow one another in an iterative process. Measures of performance reveal where improvement is necessary, and therefore what may need to be done, which is then assessed through measurement. Performance management is a broader issue than performance measurement – it is a philosophy that is supported by performance measurement.

The balanced scorecard approach

Generally, operations performance measures have been broadening in their scope. It is now widely accepted that the scope of measurement should, at some level, include external as well as internal, long-term as well as short-term, and 'soft' as well as 'hard' measures. The best-known manifestation of this trend is the 'balanced scorecard' approach taken by Kaplan and Norton. *'The balanced scorecard retains traditional financial measures. But financial measures tell the story of past events, an adequate story only for industrial age companies for which investments in long-term capabilities and customer relationships were not critical for success. These financial measures are inadequate, however, for guiding and evaluating the journey that information-age companies must make to create future value through investment in customers, suppliers, employees, processes, technology, and innovation.'*[3]

As well as including financial measures of performance, in the same way as traditional performance measurement systems, the balanced scorecard approach also attempts to provide the important information that is required to allow the overall strategy of an organisation to be reflected adequately in specific performance measures. In addition to financial measures of performance, it also includes more operational measures of customer satisfaction, internal processes, innovation and other improvement activities. In doing so, it measures the factors behind financial performance that are seen as the key drivers of future financial success. In particular, it is argued, that a balanced range of measures enables managers to address the following questions (see Figure 12.3):

- How do we look to our shareholders (financial perspective)?
- What must we excel at (internal process perspective)?
- How do our customers see us (the customer perspective)?
- How can we continue to improve and build capabilities (the learning and growth perspective)?

The balanced scorecard attempts to bring together the elements that reflect a business's strategic position, including product or service quality measures, product and service development times, customer complaints, labour productivity, and so on. At the same time, it attempts to avoid performance reporting becoming unwieldy by restricting the number of measures and focusing especially on those seen to be essential. The advantages of the approach are that it presents an overall picture of the organisation's performance in a single report, and by being comprehensive in the measures of performance it uses, encourages companies to take decisions in the interests of the whole organisation, rather than sub-optimising around narrow measures. Developing a balanced scorecard is a complex process and is now the subject of considerable debate. One of the key questions that has to be considered is how should specific measures of performance be designed? Inadequately designed performance measures can result in dysfunctional behaviour, so teams of managers are often used to develop a scorecard that reflects their organisation's specific needs.

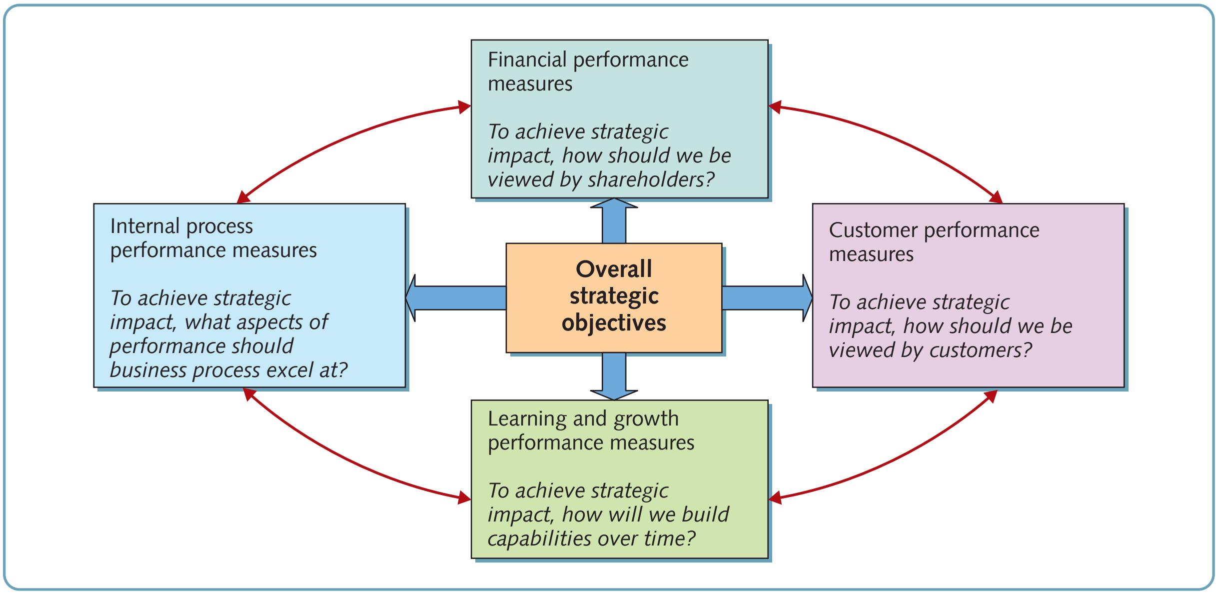


Figure 12.3 **The measures used in the balanced scorecard**

Setting target performance

A performance measure means relatively little until it is compared against some kind of target. Knowing that only one document in five hundred that is sent out to customers contains an error tells us relatively little, unless we know whether this is better or worse than we were achieving previously, and whether it is better or worse than other similar operations (especially competitors) are achieving. Setting performance targets transforms performance measures into performance 'judgements'. Several approaches to setting targets can be used, including the following:

- *Historically based targets* – targets that compare current against previous performance.
- *Strategic targets* – targets set to reflect the level of performance that is regarded as appropriate to achieve strategic objectives.
- *External performance-based targets* – targets set to reflect the performance that is achieved by similar, or competitor, external operations.
- *Absolute performance targets* – targets based on the theoretical upper limit of performance.

One of the problems in setting targets is that different targets can give very different messages regarding the improvement being achieved. So, for example, in Figure 12.4, one of an operation's performance measures is 'delivery' (in this case defined as the proportion of orders delivered on time). The performance for one month has been measured at 83 per cent, but any judgement regarding performance will be dependent on the performance targets. Using a *historical* target, when compared to last year's performance of 60 per cent, this month's performance of 83 per cent is good. But, if the operation's *strategy* calls for a 95 per cent delivery performance, the actual performance of 83 per cent looks decidedly poor. The company may also be concerned with how they perform against *competitors'* performance. If competitors are currently averaging delivery performances of around 80 per cent, the company's performance looks rather good. Finally, the more ambitious managers within the company may wish to at least try to seek perfection. Why not, they argue, use an *absolute* performance standard of 100 per cent delivery on time? Against this standard the company's actual 83 per cent again looks disappointing.

OPERATIONS PRINCIPLE
Performance measures only have meaning when compared against performance targets.

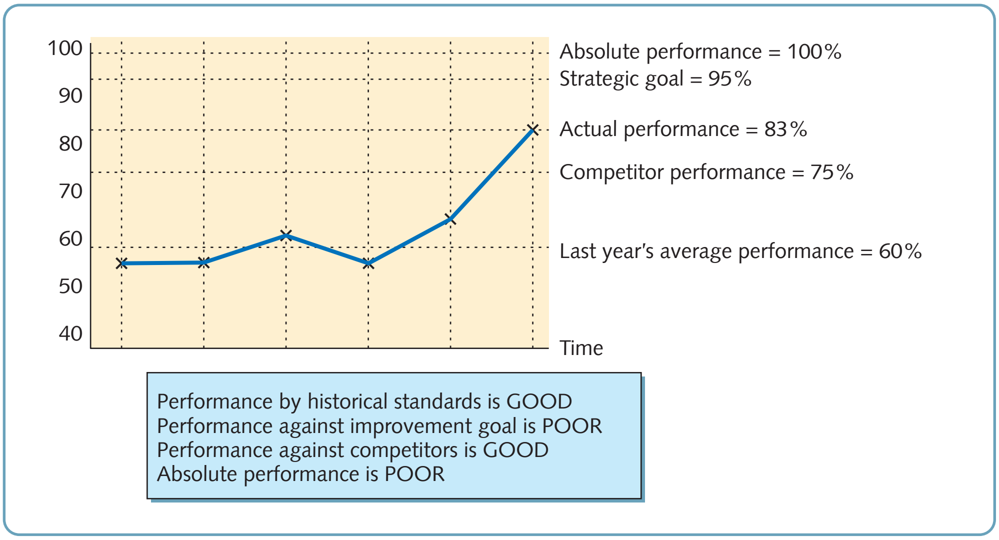


Figure 12.4 **Different standards of comparison give different messages**

Benchmarking

Benchmarking is 'the process of learning from others' and involves comparing one's own performance or methods against other comparable operations. It is a broader issue than setting performance targets, and includes investigating other organisations' operations practice in order to derive ideas that could contribute to performance improvement. Its rationale is based on the idea that (a) problems in managing processes are almost certainly shared by processes elsewhere, and (b) that there is probably another operation somewhere that has developed a better way of doing things. For example, a bank might learn some things from a supermarket about how it could cope with demand fluctuations during the day. Benchmarking is essentially about stimulating creativity in improvement practice.

> OPERATIONS PRINCIPLE
> *Improvement is aided by contextualising processes and operations.*

Case example

Triumph motorcycles resurrected through benchmarking[4]

Triumph motorcycles once built the coolest bikes in the world. In the classic prisoner-of-war film, *The Great Escape*, Steve McQueen memorably jumped across the wire on a Triumph motorcycle. In the 1960s its larger motorcycles were selling well, both in the UK and the US. But competition was catching up on the company and, just like the UK auto industry, Triumph declined from the 1970s onwards as better-designed and better-produced (mainly Japanese) products started to dominate the market. Within a few years the company had gone into receivership, and a property developer, John Bloor, bought the rights to the Triumph name relatively cheaply. He believed that there was a future for the company, yet

OLI SCARFF/AFP/Getty Images

he did not restart production immediately. Instead he spent years rethinking how the company's operations could be designed and run to compete in the modern motorcycle market. With his new team of managers, he went on an in-depth benchmarking study tour of Japan to analyse the production methods of those competitors that had driven the original Triumph into insolvency. '*We learned a lot*', says Nick Bloor, John's son, who now runs the company. It soon became clear to the management team that the original old factory in the West Midlands of the UK was not up to the task of producing world-class products. It was demolished and a new plant built in the UK that utilised the modern equipment and production methods learned on the Japanese visits. Now the company's plants in the UK and Thailand produce record numbers of bikes with styling that reflects the original bike's heritage, but with standards of engineering and reliability that match the operations that it learned from.

Types of benchmarking

There are many different types of benchmarking (which are not necessarily mutually exclusive), some of which are listed here:

- *Internal benchmarking* is a comparison between operations or parts of operations that are within the same total organisation. For example, a large motor vehicle manufacturer with several factories might choose to benchmark each factory against the others.
- *External benchmarking* is a comparison between an operation and other operations that are part of a different organisation.
- *Non-competitive benchmarking* is benchmarking against external organisations that do not compete directly in the same markets.
- *Competitive benchmarking* is a direct comparison between competitors in the same, or similar, markets.
- *Performance benchmarking* is a comparison between the levels of achieved performance in different operations. For example, an operation might compare its own performance in terms of some or all of the principal performance objectives – quality, speed, dependability, flexibility and cost – against other organisations' performance in the same dimensions.
- *Practice benchmarking* is a comparison between an organisation's operations practices, or way of doing things, and those adopted by another operation. For example, a large retail store might compare its systems and procedures for controlling stock levels with those used by another department store.

Benchmarking as an improvement tool

Although benchmarking has become popular, some businesses have failed to derive maximum benefit from it. Partly this may be because there are some misunderstandings as to what benchmarking actually entails. First, it is not a 'one-off' project; it is best practised as a continuous process of comparison. Second, it does not provide 'solutions' – rather it provides ideas and information that can lead to solutions. Third, it does not involve simply copying or imitating other operations; it is a process of learning and adapting in a pragmatic manner. Fourth, it means devoting resources to the activity. Benchmarking cannot be done without some investment, but this does not necessarily mean allocating exclusive responsibility to a set of highly paid managers. In fact, there can be advantages in organising staff at all levels to investigate and collate information from benchmarking targets. There are also some basic rules about how benchmarking can be organised:

- A prerequisite for benchmarking success is to understand thoroughly your own processes. Without this it is difficult to compare your processes against those of other companies.

- Look at the information that is available in the public domain. Published accounts, journals, conferences and professional associations can all provide information that is useful for benchmarking purposes.
- Do not discard information because it seems irrelevant. Small pieces of information only make sense in the context of other pieces of information that may emerge subsequently.
- Be sensitive in asking for information from other companies. Don't ask any questions that you would not like to be asked yourself.

Deploying external ideas

Most of the literature that deals with improvement focuses on the generation, development and deployment of improvement ideas that originated within, rather from outside, the organisation. Yet, to ignore the improvements that other companies are deploying is to ignore a potentially huge source of innovation. Whether they are competitors, suppliers, customers or simply other firms with similar challenges, firms in the wider external business environment can provide solutions to internal problems. The discussion on benchmarking is clearly related to the idea of finding inspiration from outside the organisation. But some commentators on innovation go further and argue that (legally) 'copying' from outsiders can be an effective, if underused, approach to improvement. In his book, *Copycats: How smart companies use imitation to gain a strategic edge*,[5] Oded Shenkar claims that although to argue 'imitation can be strategic seems almost blasphemous in the current scholarly climate', it can 'be strategic and should be part of the strategic repertoire of any agile firm'. In fact, 'imitation can be a differentiating factor and has the potential to deliver unique value'. He cites Apple, making the point that the iPod was not the first digital-music player; nor was the iPhone the first smartphone or the iPad the first tablet. To some extent Apple imitated ideas found in others' products but solved the technical problems, established an appropriate supply chain operating model, and made the products far more appealing. Similarly, Ray Kroc, who took McDonald's to worldwide success, copied White Castle, inventor of the fast-food burger joint. And Ireland's Ryanair imitated the business model originally developed in the US by Southwest Airlines.

Shenkar identifies three 'strategic types' of imitators:[6]

- *The pioneer importer*: An imitator that is the pioneer in another place (another country, industry or product market). This is what Ryanair did in Europe when it imported the Southwest model. Pioneer importer imitators may actually be able to move relatively slowly, especially if the original innovator (or other imitators) is unlikely to compete directly in the same market.
- *The fast second*: A rapid mover arriving quickly after an innovator or pioneer, but before it has had an opportunity to establish an unassailable lead, and before other potentially rival imitators take a large share of the market. This strategy basically lets the pioneer take much of the risk of innovation in the hope that the follower can learn from the pioneer's experience.
- *The come from behind*: A late entrant or adopter that has deliberately delayed adopting a new idea, maybe because of legal reasons, or because it wants to be more certain that the idea will be acceptable. When it does adopt the idea, it may rely on differentiating itself from the original pioneers. Samsung did this with its chip-making business, by using its manufacturing capability and knowledge to halve the time it takes to build a semiconductor plant. It then established a lead over competitors by exploiting its strength in key technical, production and quality skills.

Case example

Formula 1 transfers its ideas[7]

Some of the best improvement ideas can come from industries well away from where they may be applied. As driving jobs go, there could be no bigger difference than between Formula 1 racing drivers weaving their way through some of the fastest competitors in the world and a supermarket truck driver quietly delivering beans, beer and bacon to distribution centres and stores. But they have more in common than one would suspect. Both Formula 1 and truck drivers want to save fuel, either to reduce pit stops (Formula 1) or keep delivery costs down (heavy goods vehicles). And although grocery deliveries in the suburbs do not seem as thrilling as racing round the track at Monza, the computer-assisted simulation programs developed by the Williams Formula 1 team are being deployed to help Sainsbury's (a British supermarket group) drivers develop the driving skills that could potentially cut its fuel bill by up to 30 per cent. The simulator technology, which allows realistic advanced training to be conducted in a controlled environment, was developed originally for the advanced training of Formula 1 drivers and was developed and extended at the Williams Technology Centre in Qatar. It can now train drivers to a high level of professional driving skills and road safety applications. Because Formula 1 is well recognised as an innovative technology incubator, it makes sense for a wide range of seemingly different operations to look to motor racing for new and emerging technologies. Sainsbury's mission to reduce its energy consumption and enhance the skills and safety of those supporting its crucial logistics operation is a typical example. Its energy-related improvement programmes tackle energy supply (for example wind, solar and geothermal energy) as well as energy consumption (for example switching to LED lighting, CO_2 refrigeration, etc.). Similarly, during the COVID-19 pandemic, Formula 1's 'Project Pitlane' effort developed continuous positive airway pressure (CPAP) devices, which could help coronavirus patients with lung infections to breathe more easily.

Shahjehan/Shutterstock

Assess the gap between actual and target performance

A comparison of actual and target performance should guide the relative priorities for improvement. An important aspect of performance does not require immediate priority for improvement if current performance is already significantly better than target performance. Similarly, because some aspect of performance is relatively poor does not mean that it should be improved immediately if it exceeds its target performance. In fact, both the relative importance of the various performance measures, and their performance against target, need to be brought together in order to prioritise for improvement. One way of doing this is through the importance–performance matrix.

The importance–performance matrix

As its name implies, the importance–performance matrix positions each aspect of performance on a matrix according to its scores or ratings on how important each aspect of relative performance is, and what performance it is currently achieving. Figure 12.5 shows an importance–performance matrix divided into zones of improvement priority. The first zone boundary is the 'lower bound of acceptability', shown as line AB in Figure 12.5. This is the boundary between acceptable and unacceptable current performance. When some aspect of performance is rated as relatively unimportant, this boundary will be low. Most operations are prepared to tolerate lower performance

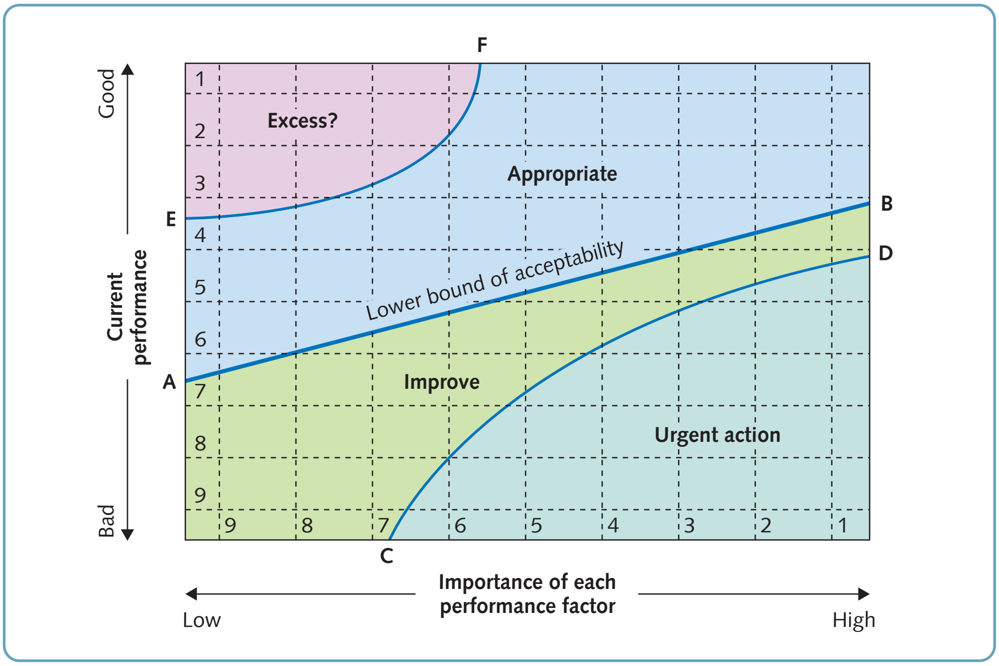


Figure 12.5 **Priority zones in the importance–performance matrix**

for relatively unimportant performance factors. However, for performance factors that are rated more important, they will be markedly less sanguine at poor or mediocre levels of current performance. Below this minimum bound of acceptability (AB) there is clearly a need for improvement; above this line there is no immediate urgency for any improvement. However, not all factors of performance that fall below the minimum line will be seen as having the same degree of improvement priority. A boundary approximately represented by line CD represents a distinction between an urgent priority zone and a less urgent improvement zone. Similarly, above the line AB, not all competitive factors are regarded as having the same priority. The line EF can be seen as the approximate boundary between performance levels that are regarded as 'good' or 'appropriate' on one hand and those regarded as 'too good' or 'excess' on the other. Segregating the matrix in this way results in four zones that imply very different priorities:

1. ***The 'appropriate' zone*** – performance factors in this area lie above the lower bound of acceptability and so should be considered satisfactory.
2. ***The 'improve' zone*** – lying below the lower bound of acceptability, any performance factors in this zone must be candidates for improvement.
3. ***The 'urgent action' zone*** – these performance factors are important to customers but current performance is unacceptable. They must be considered as candidates for immediate improvement.
4. ***The 'excess?' zone*** – performance factors in this area are 'high performing', but are not particularly important. The question must be asked, therefore, whether the resources devoted to achieving such a performance could be better used elsewhere.

The sandcone theory

As well as approaches that base improvement priority given on an operation's specific circumstances, some authorities believe that there is also a generic 'best' sequence of improvement. The best-known theory is called *the sandcone theory*,[8] so called because the sand is analogous

Worked example

EXL Laboratories

EXL Laboratories is a subsidiary of an electronics company. It carries out research and development as well as technical problem-solving work for a wide range of companies. It is particularly keen to improve the level of service that it gives to its customers. However, it needs to decide which aspects of its performance to improve first. It has devised a list of the most important aspects of its service:

- *The quality of its technical solutions* – the appropriateness perceived by customers.
- *The quality of its communications with customers* – the frequency and usefulness of information.
- *The quality of post-project documentation* – the usefulness of the documentation that goes with the final report.
- *Delivery speed* – the time between customer request and the delivery of the final report.
- *Delivery dependability* – the ability to deliver on the promised date.
- *Delivery flexibility* – the ability to deliver the report on a revised date.
- *Specification flexibility* – the ability to change the nature of the investigation.
- *Price* – the total charge to the customer.

EXL assigns a rating to each of these performance factors, both for their relative importance and their current performance, as shown in Figure 12.6. In this case, EXL have used a 1–9 scale, where 1 is 'very important', or 'very good'. Any type of scale can be used.

EXL Laboratories plotted the relative importance and current performance ratings it had given to each of its performance factors on an importance–performance matrix. This is shown in Figure 12.7. It shows that the most important aspect of performance – the ability to deliver sound technical solutions to its customers – falls comfortably within the appropriate zone. Specification flexibility and delivery flexibility are also in the appropriate zone, although only just. Both delivery speed and delivery dependability seem to be in need of improvement as each is below the minimum level of acceptability for their respective importance positions. However, two competitive factors, communications and cost/price, are clearly in need of immediate improvement. These two factors should therefore be assigned the most urgent priority for improvement. The matrix also indicates that the company's post-project documentation could almost be regarded as 'too good'.

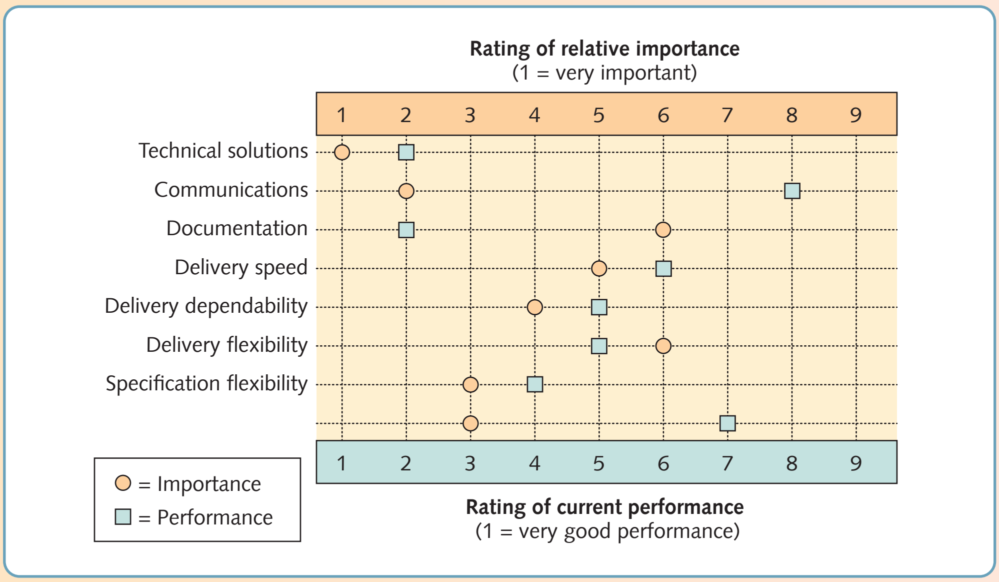


Figure 12.6 Rating 'importance' and 'current performance' for EXL Laboratories

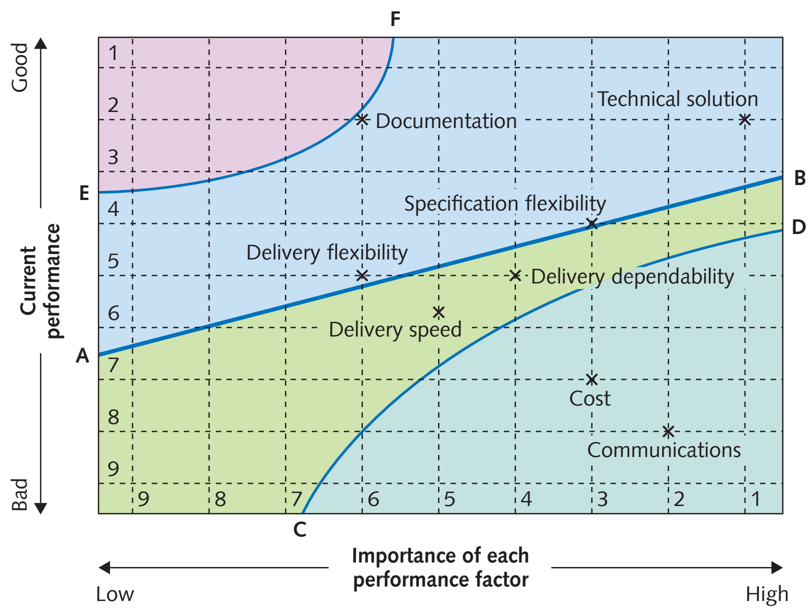


Figure 12.7 The importance–performance matrix for EXL Laboratories

to management effort and resources. Building a stable sandcone needs a stable foundation of quality, upon which one can build layers of dependability, speed, flexibility and cost (see Figure 12.8). Building up improvement is thus a cumulative process, not a sequential one. Moving on to the second priority for improvement does not mean dropping the first, and so on. According to the sandcone theory; the first priority should be *quality*, since this is a precondition to all lasting improvement. Only when the operation has reached a minimally acceptable level in quality should it then tackle the next issue, that of internal *dependability*. Importantly though, moving on to include dependability in the improvement process will actually require further improvement in quality. Once a critical level of dependability is reached

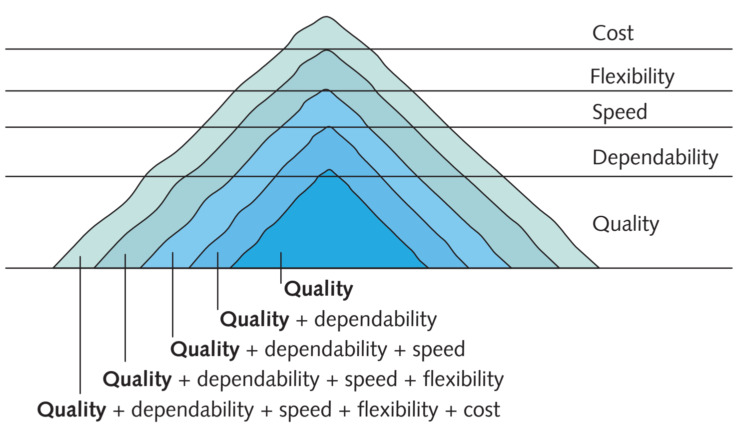


Figure 12.8 The sandcone model of improvement; cost reduction relies on a cumulative foundation of improvement in the other performance objectives

(enough to provide some stability to the operation), the next stage is to improve the *speed* of internal throughput. But again, only while continuing to improve quality and dependability further. Soon it will become evident that the most effective way to improve speed is through improvements in response *flexibility* – that is, changing things within the operation faster. Again, including flexibility in the improvement process should not divert attention from continuing to work further on quality, dependability and speed. Only now, according to the sandcone theory, should *cost* be tackled head on.

12.3 Diagnostic question: *What is the most appropriate improvement path?*

Once the priority of improvement has been determined, an operation must consider the approach or path it wishes to take to reach its improvement goals. Two paths represent different, and to some extent opposing, philosophies: *breakthrough improvement* and *continuous improvement*. Although they represent different philosophies of improvement, they are not mutually exclusive. Few operations cannot benefit from improving their operations performance on a continuous basis, and few operations would reject investing in a major improvement breakthrough leap in performance if it represented good value. For most operations, both approaches are relevant to some extent, although possibly at different points in time. But to understand how and when each approach is appropriate, one must understand their underlying philosophies.

> OPERATIONS PRINCIPLE
> ***Breakthrough and continuous improvement are not mutually exclusive.***

Breakthrough improvement

Breakthrough (or 'innovation'-based) improvement assumes that the main vehicle of improvement is major and dramatic change in the way the operation works – for example, the total reorganisation of an operation's process structure, or the introduction of a fully integrated information system. The impact of these improvements represents a step change in practice (and hopefully performance). Such improvements can be expensive, often disrupting the ongoing workings of the operation, and frequently involving changes in the product/service or process technology. Moreover, the actual level of improvement is not guaranteed. In fact, some improvement specialists argue that the greater the intended step change in performance, the greater the risk that the actual increase in performance will disappoint.

The business process re-engineering approach

Typical of the radical breakthrough way of tackling improvement is the business process re-engineering (BPR) approach. It is a blend of a number of ideas such as fast throughput, waste elimination through process flow charting, customer-focused operations, and so on. But it was the potential of information technologies to enable the fundamental redesign of processes that acted as the catalyst in bringing these ideas together. BPR has been described as the fundamental rethinking and radical redesign of business processes to achieve dramatic improvements in critical contemporary measures of performance, such as cost, quality, service and speed.[9]

Underlying the BPR approach is the belief that operations should be organised around the total process that adds value for customers, rather than the functions or activities that perform the various stages of the value-adding activity. (Figure 1.10 in Chapter 1 illustrates this idea.) The core of BPR is a redefinition of the processes within a total operation, to reflect the business

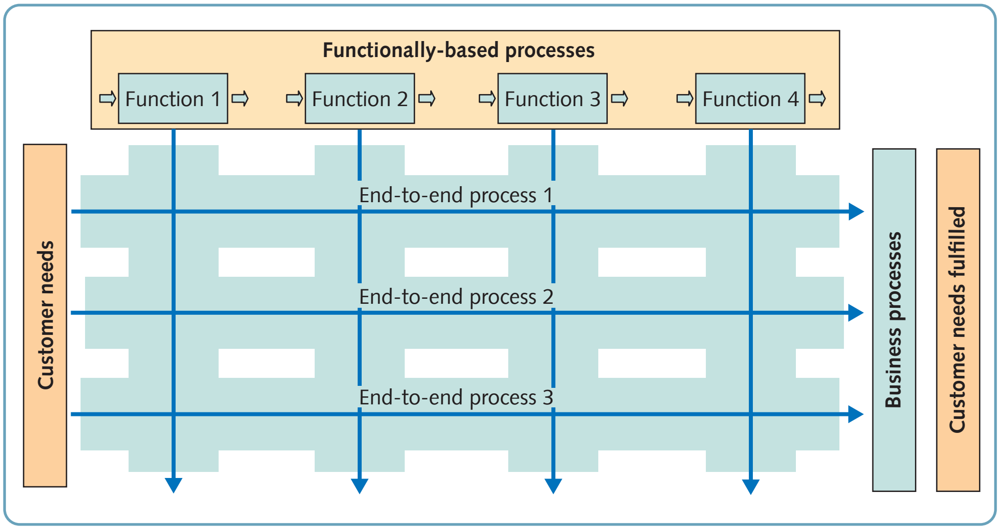


Figure 12.9 **BPR advocates reorganising (re-engineering) processes to reflect the natural processes that fulfil customer needs**

processes that satisfy customer needs (see Figure 12.9). The main principles of BPR have been summarised as follows:[10]

- Rethink business processes in a cross-functional manner that organises work around the natural flow of information (or materials or customers). This means organising around outcomes of a process rather than the tasks that go into it.
- Strive for dramatic improvements in the performance by radically rethinking and redesigning the process.
- Have those who use the output from a process perform the process. Check to see if all internal customers can be their own supplier, rather than depending on another function in the business to supply them (which takes longer and separates out the stages in the process).
- Put decision points where the work is performed. Do not separate those who do the work from those who control and manage the work. Control and action are just one more type of supplier–customer relationship that can be merged.

Continuous improvement

Continuous improvement, as the name implies, adopts an approach to improving performance that assumes a never-ending series of small incremental improvement steps – for example, modifying the way a product is fixed to a machine to reduce changeover time, or simplifying the question sequence when taking a hotel reservation. While there is no guarantee that other steps will follow such small steps towards better performance, the whole philosophy of continuous improvement attempts to ensure that there will be. It is also known as *kaizen*, after the Japanese word that translates roughly as 'improvement', meaning improvement in personal, home, social and work life. When applied to operations, *kaizen* means continuing improvement involving everyone at every level of the organisation.

Continuous improvement is not concerned with promoting small improvements *per se*, but it does view small improvements as having one significant advantage over large ones – they can be followed relatively painlessly by others. It is not the *rate* of improvement that is important; it is the *momentum* of improvement. It does not matter if successive improvements are small; what does matter is that every month (or week, or quarter, or whatever period is appropriate)

Case example

Kaizen at Amazon[11]

At one point in Amazon's development of its fulfilment centres it believed that most issues could be solved with technology. It changed its mind as it learned that engaging its front-line staff in continuous improvement could be at least equally effective. For example, the company had been attempting to automate a large part of its fulfilment centres. However, the automation was devised for books and it did not work well on the other goods that Amazon was introducing, such as shoes. So, when the shoebox reached the mechanism in the automated system that was intended to bring the shoes to the packing line, they were hurled out of the box. Given its experience, Amazon adopted an 'autonomation' approach, using people for complex tasks and using machines to support them. In another *kaizen* improvement, the time to scan products being put onto shelves in one of the fulfilment centres, called the 'stow line', was taking longer than anticipated. Each person on the line had a trolley full of products to stow on the shelves, and a scanner. The products and the corresponding shelf number had to be scanned so that the computer knew where each product was located. The standard time target for this task was 20 minutes per trolley. But when one of the company's senior managers tried to perform this task, it took him 45 minutes. One of the reasons was that he had to scan some items four times before the scanner recognised them. It was obvious that, at least in part, rather than simply being incompetent, his performance was affected by an abnormality – the bad performance of the scanner. After analysing all the deviations from expected performance that had been reported by the staff, and looking for their root causes, they found that managers were unclear as to how battery life affected scanner performance. In fact, there were several hours of low productivity because of the low battery charge at the end of each scanner's operating time, and there was no satisfactory process in place to check and reload the scanner batteries. This root-cause analysis helped them to put a new process in place to monitor and load the scanners to avoid low-charge periods.

some kind of improvement has actually taken place. Continuous improvement does not always come naturally. There are specific abilities, behaviours and actions that need to be consciously developed if continuous improvement is to sustain over the long term.

Improvement-cycle models

An important element of continuous improvement is the idea that improvement can be represented by a never-ending process of repeatedly questioning and re-questioning the detailed working of a process. This is usually summarised by the idea of the *improvement cycle*, of which there are many, some proprietary models owned by consultancy companies. Two of the more generally used models are the PDCA cycle (sometimes called the Deming cycle, named after the famous quality 'guru', W.E. Deming), and the DMAIC cycle (made popular by the Six-Sigma approach to improvement – see later).

> OPERATIONS PRINCIPLE
> ***Continuous improvement necessarily implies a never-ending cycle of analysis and action.***

The PDCA (or PDSA) cycle

The PDCA cycle model is shown in Figure 12.10(a). It starts with the P (for plan) stage, which involves an examination of the current method or the problem area being studied. This involves collecting and analysing data, so as to formulate a plan of action that is intended to improve performance. (Some of the techniques used to collect and analyse data are explained later.) The next step is the D (for do) stage. This is the implementation stage, during which the plan is tried out in the operation. This stage may itself involve a mini-PDCA cycle as the problems of implementation are resolved. Next comes the C (for check) stage where the newly implemented solution is evaluated to see whether it has resulted in the expected improvement. Some versions of this idea use the term 'study' instead of 'check' and call the idea the 'PDSA' cycle,

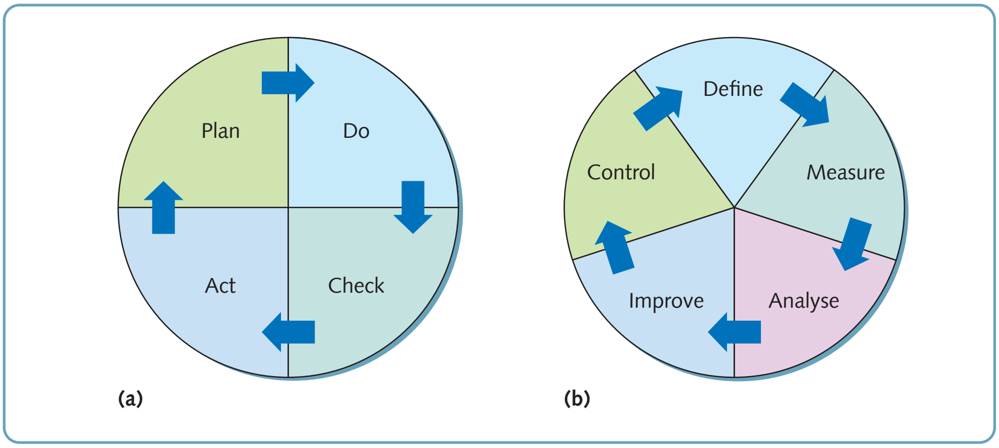


Figure 12.10 **(a) the plan-do-check-act, or 'Deming' improvement cycle and (b) the define-measure-analyse-improve-control, or DMAIC Six Sigma improvement cycle**

but the idea is basically the same. Finally, at least for this cycle, comes the A (for act) stage. During this stage the change is consolidated or standardised, if it has been successful. Alternatively, if the change has not been successful, the lessons learned from the 'trial' are formalised before the cycle starts again. You may also find this cycle called the Deming cycle, Deming wheel or Shewhart cycle.

The DMAIC cycle

In some ways, this cycle is more intuitively obvious than the PDCA cycle in so much as it follows a more 'experimental' approach. The DMAIC cycle (see Figure 12.10(b)) starts with defining the problem or problems, partly to understand the scope of what needs to be done and partly to define exactly the requirements of the process improvement. Often at this stage, a formal goal or target for the improvement is set. After definition comes the measurement stage, important because the Six Sigma approach emphasises the importance of working with hard evidence rather than opinion. It involves validating the problem (to make sure it is really worth solving), using data to refine the problem and measuring exactly what is happening. The analysis stage can be seen as an opportunity to develop hypotheses about what the root causes of the problem really are. Such hypotheses are validated (or not) by the analysis and the main root causes of the problem identified. Once the causes of the problem are identified, work can begin on improving the process. Ideas are developed to remove the root causes of problems, solutions are tested and those solutions that seem to work are implemented, formalised and results measured. The improved process then needs to be continually monitored and controlled, to check that the improved level of performance is sustained. The cycle then starts again, defining the problems that are preventing further improvement.

The last point in both the PDCA and DMAIC cycles is the most important – *'the cycle starts again'*. It is only by accepting that in a continuous improvement philosophy these cycles quite literally never stop, that improvement becomes part of every person's job.

The differences between breakthrough and continuous improvement

Breakthrough improvement places a high value on creative solutions, and encourages free thinking and individualism. It is a radical philosophy in so much as it fosters an approach to improvement that does not accept many constraints on what is possible. 'Starting with a clean sheet of paper', 'going back to first principles' and 'completely rethinking the system'

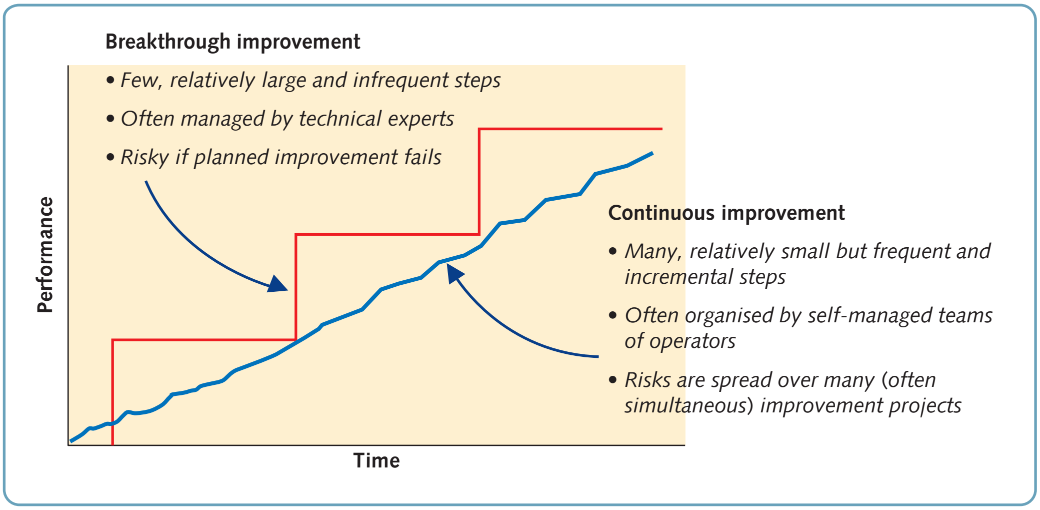


Figure 12.11 The differences between the breakthrough and continuous approaches to improvement

are all typical breakthrough improvement principles. Continuous improvement, on the other hand, is less ambitious, at least in the short term. It stresses adaptability, teamwork and attention to detail. It is not radical; rather it builds upon the wealth of accumulated experience within the operation itself, often relying primarily on the people who operate the system to improve it. One analogy used to explain this difference is the sprint versus marathon. Breakthrough improvement is a series of explosive and impressive sprints. Continuous improvement, like marathon running, does not require the expertise and prowess required for sprinting; but it does require that the runner (or operations manager) keeps on going. Yet notwithstanding these differences, it is possible to use both approaches. Large and dramatic improvements can be implemented as and when they seem to promise significant improvement steps, but between such occasions the operation can continue making its quiet and less spectacular *kaizen* improvements. Figure 12.11 shows some of the differences between the two approaches.

OPERATIONS PRINCIPLE
Breakthrough improvement necessarily implies radical and/or extensive change.

Exploitation or exploration

A closely related distinction to that between continuous and breakthrough improvement, is the one that management theorists draw between what they call 'exploitation' versus 'exploration'. Exploitation is the activity of enhancing processes (and products) that already exist within a firm. The focus of exploitation is on creating efficiencies rather than radically changing resources or processes. Its emphasis is on tight control of the improvement process, standardising processes, clear organisational structures and organisational stability. The benefits from exploitation tend to be relatively immediate, incremental and predictable. They also are likely to be better understood by the firm and fit into its existing strategic framework. Exploration, by contrast, is concerned with the exploration of new possibilities. It is associated with searching for and recognising new mindsets and ways of doing things. It involves experimentation, taking risks, simulation of possible consequences, flexibility and innovation. The benefits from exploration are principally long term but can be relatively difficult to predict. Moreover, any benefits or discoveries that might come may be so different from what the firm is familiar with that it may not find it easy to take advantage of them.

Organisational 'ambidexterity'

It is clear that the organisational skills and capabilities needed to be successful at exploitation are likely to be very different from those that are needed for the radical exploration of new ideas. Indeed, the two views of improvement may actively conflict. A focus on thoroughly exploring for totally novel choices may consume managerial time, effort and the financial resources that would otherwise be used for refining existing ways of doing things, thereby reducing the effectiveness of improving existing processes. Conversely, if existing processes are improved over time, there may be less motivation to experiment with new ideas. So, although both exploitation and exploration can be beneficial, they may compete both for resources and for management attention. This is where the concept of 'organisational ambidexterity' becomes important. Organisational ambidexterity means the ability of a firm to both exploit and explore as it seeks to improve; to be able to compete in mature markets where efficiency is important, by improving existing resources and processes, while also competing in new technologies and/or markets where novelty, innovation and experimentation are required.

The Six Sigma approach to organising improvement

One approach to improvement that combines breakthrough and continuous philosophies is Six Sigma. Although technically the 'Six Sigma' name derives from statistical process control (SPC), and more specifically the concept of process capability, it has now come to mean a much broader approach to improvement. The following description gives a sense of its modern usage: '*Six Sigma is a comprehensive and flexible system for achieving, sustaining and maximizing business success. Six Sigma is uniquely driven by close understanding of customer needs, disciplined use of facts, data, and statistical analysis, and diligent attention to managing, improving, and reinventing business processes*.'[12]

The Six Sigma concept, therefore, includes many of the issues covered in this and other chapters of this text. For example, process design and redesign, balanced scorecard measures, continuous improvement, statistical process control, ongoing process planning and control, and so on. However, at the heart of Six Sigma lies an understanding of the negative effects of variation in all types of business process. This aversion to variation was first popularised by Motorola, the electronics company, which set its objective as 'total customer satisfaction' in the 1980s, then decided that true customer satisfaction would only be achieved when its products were delivered when promised, with no defects, with no early-life failures and no excessive failure in service. To achieve this, it focused initially on removing manufacturing defects, but it soon realised that many problems were caused by latent defects, hidden within the design of its products. The only way to eliminate these defects was to make sure that design specifications were tight (i.e. narrow tolerances) and its processes very capable.

Motorola's Six Sigma quality concept was so named because it required that the natural variation of processes ($\pm$ 3 standard deviations) should be half its specification range. In other words, the specification range of any part of a product or service should be $\pm$ 6 the standard deviation of the process. The Greek letter sigma (σ) is often used to indicate the standard deviation of a process, hence the Six Sigma label. The Six Sigma approach also used the measure of 'defects per million *opportunities*' (DPMO). This is the number of defects that the process will produce if there were one million opportunities to do so. So difficult processes with many opportunities for defects can be compared with simple processes with few opportunities for defects.

The Six Sigma approach also holds that improvement initiatives can be successful only if significant resources and training are devoted to their management. It recommends a specially trained cadre of practitioners, many of whom should be dedicated full time to improving processes as internal consultants. The terms that have become associated with this group of experts (and denote their level of expertise) are, Master Black Belt, Black Belt, and Green Belt.

Master Black Belts – are experts in the use of Six Sigma tools and techniques, as well as how such techniques can be used and implemented. They are seen as teachers who can not only guide improvement projects, but also coach and mentor Black Belts and Green Belts. Given their responsibilities, it is expected that Master Black Belts are employed full time on their improvement activities.

Black Belts – take a direct hand in organising improvement teams, and will usually have undertaken a minimum of 20–25 days training and carried out at least one major improvement project. Black Belts are expected to develop their quantitative analytical skills and also act as coaches for Green Belts. Like Master Black Belts, they are dedicated full time to improvement and, although opinions vary, some organisations recommend one Black Belt for every 100 employees.

Green Belts – work within improvement teams, possibly as team leaders. They have less training than Black Belts – typically around 10–15 days. Green Belts are not full-time positions. They have normal day-to-day process responsibilities but are expected to spend at least 20 per cent of their time on improvement projects.

Devoting such large amounts of training and time to improvement is a significant investment, especially for small companies. Nevertheless, Six Sigma proponents argue that the improvement activity is generally neglected in most operations and, if it is to be taken seriously, it deserves the significant investment implied by the Six Sigma approach. Furthermore, they argue, if operated well, Six Sigma improvement projects run by experienced practitioners can save far more than they cost.

Case example

Six Sigma at Wipro[13]

There are many companies that have benefited from Six Sigma-based improvement, but few have gone on to be able to sell the expertise that they gathered from applying it to themselves. Wipro is one of these. Wipro is a global information technology, consulting and outsourcing company with 145,000 employees serving over 900 clients in 60 countries. It provides a range of business services from 'business process outsourcing' (doing processing for other firms) to 'software development', and from 'information technology consulting' to 'cloud computing'. (Surprisingly for a global IT services giant, Wipro was actually started in 1945 in India as a vegetable oil company.) Wipro also has one of the most developed Six Sigma programmes in the IT and consulting industry, especially in its software development activities, where key challenges include reducing the data transfer time within the process, reducing the risk of failures and errors, and avoiding interruption due to network downtime. For Wipro, Six Sigma simply means a measure of quality that strives for near perfection. It means:

- having products and services that meet global standards;
- ensuring robust processes within the organisation;
- consistently meeting and exceeding customer expectations; and
- establishing a quality culture throughout the business.

Individual Six Sigma projects were selected on the basis of their probability of success and were completed relatively quickly. This gave Wipro the opportunity to assess the success and learn from any problems that had occurred. Projects were identified on the basis of the problem areas under each of the critical business processes that could adversely impact business performance. Because Wipro took a customer-focused definition of quality, Six Sigma implementation was measured in terms of progress towards what the customer finds important (and what the customer pays for). This involved improving performance through a precise quantitative understanding of the customer's requirements. Wipro says that its adoption of Six Sigma has been an unquestionable success, whether in terms of customer satisfaction, improvement in internal performance or in the improvement of shareholder value.

However, as the pioneers of Six Sigma in India, Wipro's implementation of Six Sigma was not without difficulties – and, they stress, opportunities for learning from them.

To begin with, it took time to build the required support from the higher-level managers, and to restructure the organisation to provide the infrastructure and training to establish confidence in the process. In particular, the first year of deployment was extremely difficult. Resourcing the stream of Six Sigma projects was problematic, partly because each project required different levels and types of resource. Also, the company learned not to underestimate the amount of training that would be required. To build a team of professionals and train them for various stages of Six Sigma was a difficult job. (In fact, this motivated Wipro to start its own consultancy that could train its own people.) Nevertheless, regular and timely reviews of each project proved particularly important in ensuring the success of a project and Wipro had to develop a team of experts for this purpose.

12.4 Diagnostic question: *What techniques could be used to facilitate improvement?*

OPERATIONS PRINCIPLE
Improvement can be facilitated by relatively simple analytical techniques.

All the techniques described in this text and its supplements can be regarded as 'improvement' techniques. Describing all improvement techniques that have been used in all industries would be impractical. However, some techniques are widely used, and useful for improving operations and processes generally. Here we select some techniques that either have not been described elsewhere or need to be reintroduced in their role of helping operations improvement.

Scatter diagrams

Scatter diagrams provide a quick and simple method of identifying whether there is evidence of a connection between two sets of data: for example, the time at which you set off for work every morning and how long the journey to work takes. Plotting each journey on a graph, which has departure time on one axis and journey time on the other, could give an indication of whether departure time and journey time are related, and if so, how. Scatter diagrams can be treated in a far more sophisticated manner by quantifying how strong the relationship between the sets of data is. But, however sophisticated the approach, this type of graph only identifies the existence of a relationship, not necessarily the existence of a cause–effect relationship. If the scatter diagram shows a very strong connection between the sets of data, it is important evidence of a cause–effect relationship, but not proof positive. It could be coincidence!

Worked example

Kaston Pyral Services Ltd (1)

Kaston Pyral Services Ltd. (KPS) installs and maintains environmental control, heating and air conditioning systems. It set up an improvement team to suggest ways in which it might improve its levels of customer service. The improvement team had completed its first customer satisfaction survey. The survey asked customers to score the service they received from KPS in several ways. For example, it asked customers to score services on a scale of 1–10 on promptness, friendliness, level of advice, etc. Scores were then summed to give a 'total satisfaction score' for each customer – the higher the score, the greater the satisfaction. The spread of satisfaction scores puzzled the team and they considered what factors might be causing such differences in the way their customers viewed them. Two factors were put forward to explain the differences:

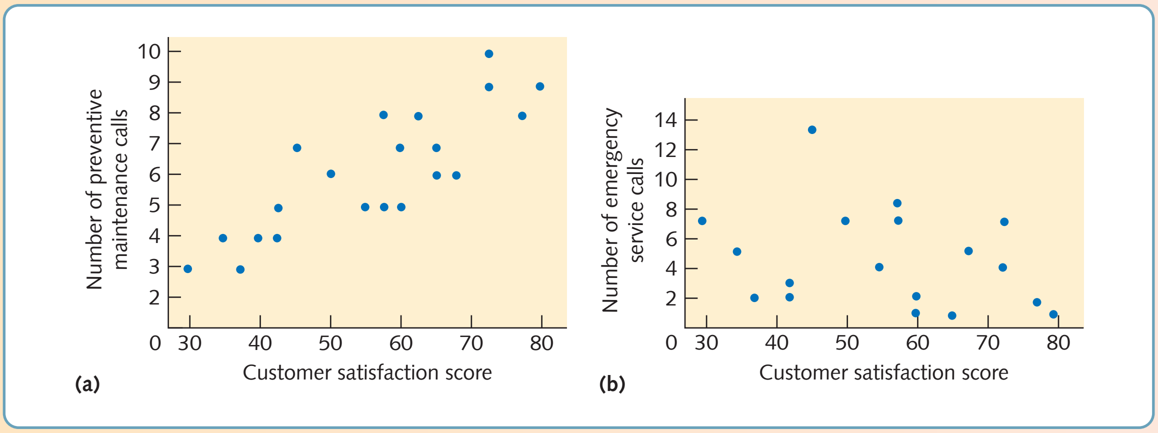


Figure 12.12 Scatter diagrams for customer satisfaction versus (a) number of preventive maintenance calls and (b) number of emergency service calls

1. the number of times in the past year the customer had received a preventive maintenance visit;
2. the number of times the customer had called for emergency service.

All this data was collected and plotted on scatter diagrams as shown in Figure 12.12. Figure 12.12(a) shows that there seems to be a clear relationship between a customer's satisfaction score and the number of times the customer was visited for regular servicing. The scatter diagram in Figure 12.12(b) is less clear. Although all customers who had very high satisfaction scores had made very few emergency calls, so had some customers with low satisfaction scores. As a result of this analysis, the team decided to survey customers' views on its emergency service.

Cause–effect diagrams

Cause–effect diagrams are a particularly effective method of helping to search for the root causes of problems. They do this by asking what, when, where, how and why questions, but also add some possible 'answers' in an explicit way. They can also be used to identify areas where further data is needed. Cause–effect diagrams (which are also known as Ishikawa diagrams) have become extensively used in improvement programmes. This is because they provide a way of structuring group brainstorming sessions. Often the structure involves identifying possible causes under the (rather old-fashioned) headings of: machinery, manpower, materials, methods and money. Yet, in practice, any categorisation that comprehensively covers all relevant possible causes could be used.

Worked example

Kaston Pyral Services Ltd (2)

The improvement team at KPS was working on a particular area that was proving a problem. Whenever service engineers were called out to perform emergency servicing for a customer, they took with them the spares and equipment that they thought would be necessary to repair the system. Although engineers could never be sure exactly what materials and equipment they would need for a job, they could guess what was likely to be needed and take a range of spares and equipment that would cover most eventualities. Too often, however, the engineers would find that they needed a spare that they had not brought with them. The cause–effect diagram for this particular problem, as drawn by the team, is shown in Figure 12.13.

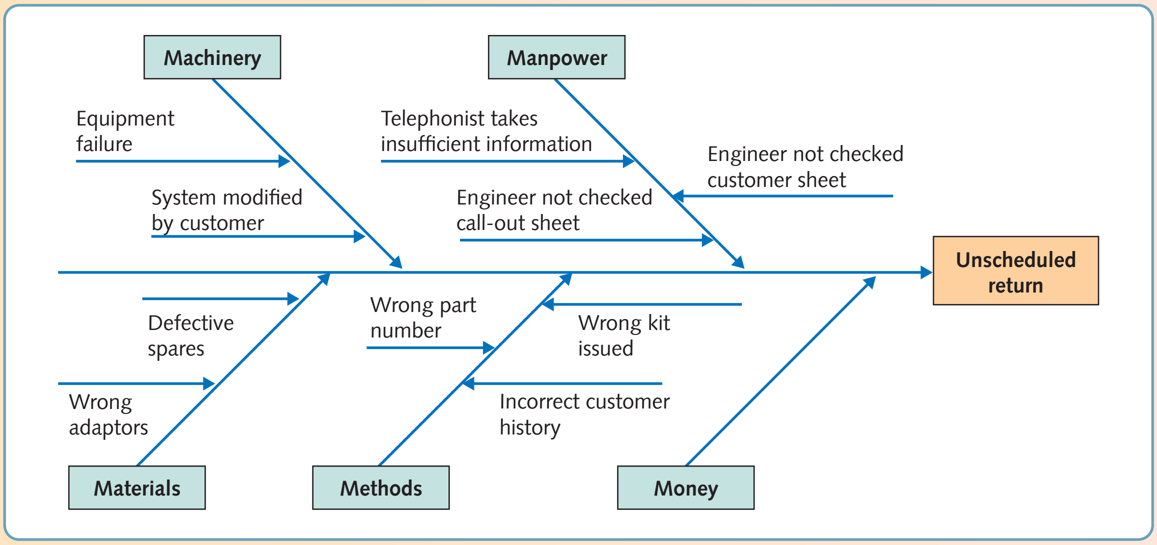


Figure 12.13 Cause–effect diagram of unscheduled returns at KPS

Pareto diagrams

In any improvement process, it is worthwhile distinguishing what is important and what is less so. The purpose of the Pareto diagram (which was first introduced in Chapter 9) is to distinguish between the 'vital few' issues and the 'trivial many'. It is a relatively straightforward technique that involves arranging items of information on the types or causes of problems into their order of importance (usually measured by 'frequency of occurrence'). This can be used to highlight areas where further decision-making will be useful. Pareto analysis is based on the phenomenon of relatively few causes explaining the majority of effects. For example, most revenue for any company is likely to come from relatively few of the company's customers. Similarly, relatively few of a doctor's patients will probably occupy most of his or her time.

Worked example

Kaston Pyral Services Ltd (3)

The KPS improvement team that was investigating unscheduled returns from emergency servicing (the issue that was described in the cause–effect diagram in Figure 12.13) examined all occasions over the previous 12 months on which an unscheduled return had been made. They categorised the reasons for unscheduled returns as follows:

1. The wrong part had been taken to a job because, although the information that the engineer received was sound, he or she had incorrectly predicted the nature of the fault.
2. The wrong part had been taken to the job because there was insufficient information given when the call was taken.
3. The wrong part had been taken to the job because the system had been modified in some way not recorded in KPS's records.
4. The wrong part had been taken to the job because the part had been incorrectly issued to the engineer by stores.

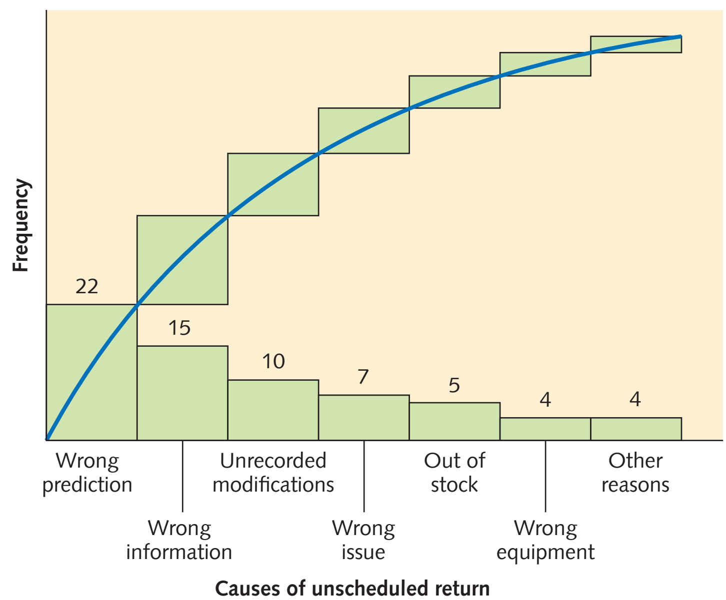


Figure 12.14 **Pareto diagram for causes of unscheduled returns**

5. No part had been taken because the relevant part was out of stock.
6. The wrong equipment had been taken for whatever reason.
7. Any other reason.

The relative frequency of occurrence of these causes is shown in Figure 12.14. About a third of all unscheduled returns were due to the first category, and more than half the returns were accounted for by the first and second categories together. It was decided that the problem could best be tackled by concentrating on how to get more information to the engineers that would enable them to predict the causes of failure accurately.

Why–why analysis

Why–why analysis starts by stating the problem and asking *why* that problem has occurred. Once the major reasons for the problem occurring have been identified, each of the major reasons is taken in turn, and again the question is asked *why* those reasons have occurred, and so on. This procedure is continued until either a cause seems sufficiently self-contained to be addressed by itself or no more answers to the question 'Why?' can be generated.

Worked example

Kaston Pyral Services Ltd (4)

The major cause of unscheduled returns at KPS was the incorrect prediction of reasons for the customer's system failure. This is stated as the 'problem' in the why–why analysis in Figure 12.15. The question is then asked, why was the failure wrongly predicted? Three answers are proposed: first, that the engineers were not trained correctly; second, that they had insufficient knowledge of the particular product installed in the customer's location; and third, that they had insufficient knowledge of the customer's particular system with its modifications. Each of these three reasons is taken in turn, and the questions are asked, why is there a lack of training, why is there a lack of product knowledge, and why is there a lack of customer knowledge? And so on.

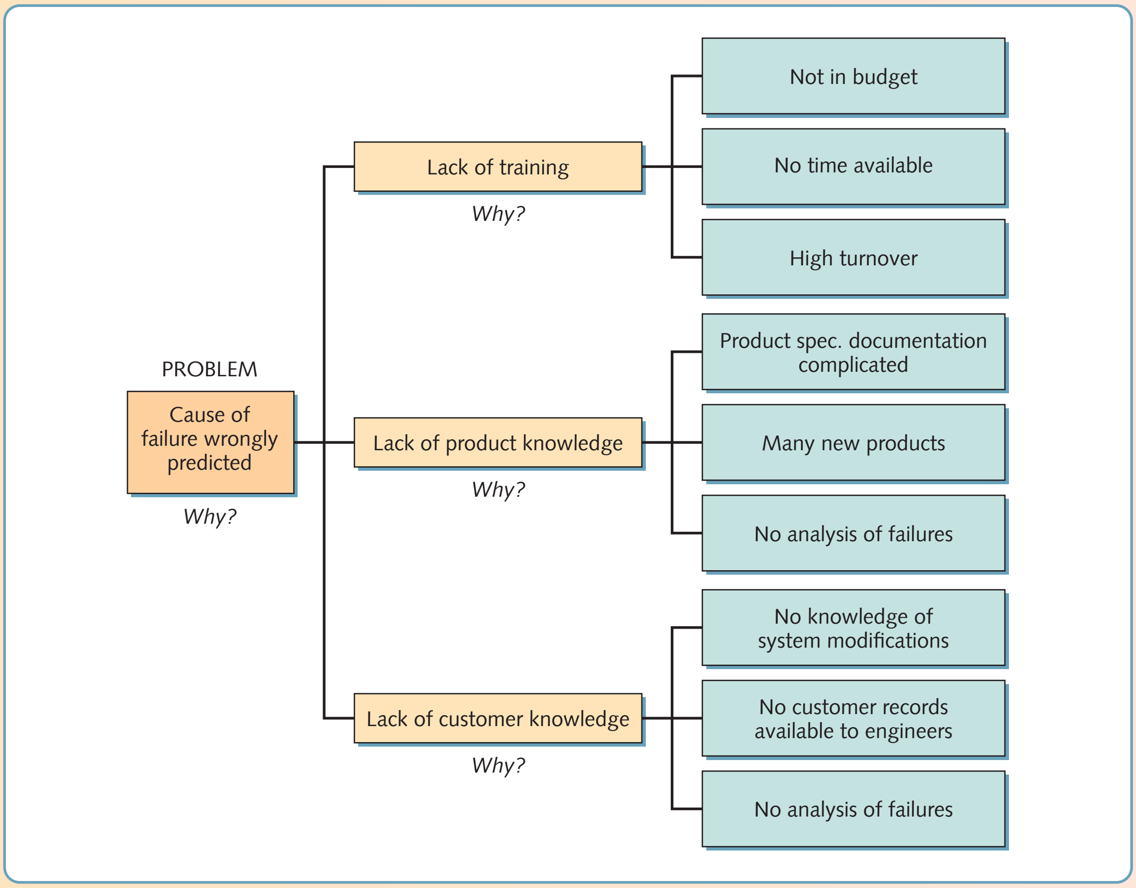


Figure 12.15 **Why–why analysis for 'failure wrongly predicted'**

12.5 Diagnostic question: *How can improvement be made to stick?*

Not all of the improvement initiatives (often launched with high expectations) will go on to fulfil their potential. Even those improvement initiatives that are successfully implemented may lose impetus over time. Sometimes this is because of managers' view of the nature of improvement; at other times it is because managers fail to manage the improvement process adequately. Nor does a successful improvement culture necessarily rely on the kinds of techniques described earlier in this chapter (although they clearly have a role). Sometimes the simplest approaches can be remarkably effective.

Avoid becoming a victim of improvement 'fashion'

Improvement has, to some extent, become a fashion industry with new ideas and concepts continually being introduced as offering a novel way to improve business performance. There is nothing intrinsically wrong with this. Fashion stimulates and refreshes through introducing novel

Case example

Simple checklists in healthcare[14]

Improvement methodologies are often associated with repetitive operations. Performing the same task repeatedly means that there are plenty of opportunities to 'get it right'. The whole idea behind continuous improvement derives from this simple idea. By contrast, operations that have to perform more difficult activities, especially those that call for expert judgement and diagnostic ability, must call for equally complex improvement approaches. Well no, according to Atul Gawande, a physician at the prestigious Johns Hopkins Hospital. Mr Gawande thinks that the very opposite is true, although medicine is advancing at an astounding rate and medical journals produce learned papers adding the results of advanced research to an ever-expanding pool of knowledge. Surgeons carry out over 200 major operations a year, unfortunately not all of them successful, but the medical profession overall does not always have a reliable method for learning from its mistakes. Atul Gawande's idea is that his and similar 'knowledge-based' professions are in danger of sinking under the weight of facts. Scientists are accumulating more and more information and professions are fragmenting into ever-narrower specialisms. Mr Gawande tells the story of Peter Pronovost, a specialist in critical care at Johns Hopkins Hospital, who tried to reduce the number of patients that were becoming infected on account of the use of intravenous central lines. There are five steps that medical teams can take to reduce the chances of contracting such infections. Initially Pronovost simply asked nurses to observe whether doctors took the five steps. What they found was that, at least a third of the time, they missed one or more of the steps. So nurses were authorised to stop doctors who had missed out any of the steps, and, as a matter of course, ask whether existing intravenous central lines should be reviewed. As a result of applying these simple checklist-style rules, the ten-day line-infection rates went down from 11 per cent to zero. In one hospital, it was calculated that, over a year, this simple method had prevented 43 infections, eight deaths and saved about $2 million. Using the same checklist approach the hospital identified and applied the method to other activities. For example, a check in which nurses asked patients about their pain levels led to untreated pain reducing from 41 per cent to 3 per cent. Similarly, the simple checklist method helped the average length of patient stay in intensive care to fall by half. When Pronovost's approach was adopted by other hospitals, within 18 months 1,500 lives and $175 million had been saved. Mr Gawande describes checklists used in this way as a 'cognitive net' – a mechanism that can help prevent experienced people from making errors due to flawed memory and attention, and ensure that teams work together Even for pilots, many of whom are rugged individualists, says Mr Gawande, it is usually the application of routine procedures that saves planes when things go wrong, rather than 'hero-pilotry' so fêted by the media. It is discipline rather than brilliance that preserves life. In fact, it is discipline that leaves room for brilliance to flourish.

OPERATIONS PRINCIPLE
The popularity of an improvement approach is not necessarily an indicator of its effectiveness.

ideas. Without it, things would stagnate. The problem lies not with new improvement ideas, but rather with some managers becoming victims of the process, where some new idea will entirely displace whatever went before. Most new ideas have something to say, but jumping from one fad to another will not only generate a backlash against any new idea, but also destroy the ability to accumulate the experience that comes from experimenting with each one.

Avoiding becoming an improvement fashion victim is not easy. It requires that those directing the improvement process take responsibility for a number of issues:

- They must take responsibility for improvement as an ongoing activity, rather than becoming champions for only one specific improvement initiative.
- They must take responsibility for understanding the underlying ideas behind each new concept. Improvement is not 'following a recipe' or 'painting by numbers'. Unless one understands *why* improvement ideas are supposed to work, it is difficult to understand *how* they can be made to work properly.

- They must take responsibility for understanding the antecedents to a 'new' improvement idea, because it helps to understand it better and to judge how appropriate it may be for one's own operation.
- They must be prepared to adapt a new idea so that it makes sense within the context of their own operation. 'One size' rarely fits all.
- They must take responsibility for the (often significant) education and learning effort that will be needed if new ideas are to be intelligently exploited.
- Above all, they must avoid the over-exaggeration and hype that many new ideas attract. Although it is sometimes tempting to exploit the motivational 'pull' of new ideas through slogans, posters and exhortations, carefully thought-out plans will always be superior in the long run, and will help avoid the inevitable backlash that follows 'overselling' a single approach.

Managing the improvement process

> OPERATIONS PRINCIPLE
> ***There is no one universal approach to improvement.***

There is no absolute prescription for the way improvement should be managed. Any improvement process should reflect the uniqueness of each operation's characteristics. What appears to be almost a guarantee of difficulty in managing improvement processes are attempts to squeeze improvement into a standard mould. Nevertheless, there are some aspects of any improvement process that appear to influence its eventual success, and should at least be debated.

Should an improvement *strategy* be defined?

Without thinking through the overall purpose and long-term goals of the improvement process it is difficult for any operation to know where it is going. Specifically, an improvement strategy should have something to say about:

- the competitive priorities of the organisation, and how the improvement process is expected to contribute to achieving increased strategic impact;
- the roles and responsibilities of the various parts of the organisation in the improvement process;
- the resources that will be available for the improvement process;
- the general approach to, and philosophy of, improvement in the organisation.

Yet, too rigid a strategy can become inappropriate if the business's competitive circumstances change, or as the operation learns through experience. But, the careful modification of improvement strategy in the light of experience is not the same as making dramatic changes in improvement strategy as new improvement fashions appear.

What degree of top-management support is required?

For most authorities, the answer is unambiguous – a significant amount. Without top-management support, improvement cannot succeed. It is the most crucial factor in almost all the studies of improvement process implementation. It also goes far beyond merely allocating senior resources to the process. 'Top-management support' usually means that senior personnel must:

- understand and believe in the link between improvement and the business's overall strategic impact;

Case example

Michelin's 'responsabilisation'[15]

Michelin, the French multinational tyre manufacturer, has developed its own approach to organising improvement. It has an attitude to empowerment that, in French, is called 'responsabilisation', which roughly translates into English as a mixture of empowerment and accountability. It is an initiative that is part of the group's efforts to streamline their organisational structures, increase responsiveness and efficiency, and encourage faster improvement and decision-making. Team empowerment is seen as an essential part of this aim. Not only does it foster initiative and dialogue, it also enables decisions to be made close to operations and customers. The aim is to develop trust-based relationships throughout the group, which encourage all employees to take part in improvement. It involves empowering front-line teams to organise themselves and find the right solutions to meet given objectives. Michelin believes that managers can then take on the role of advisors who train their teams and develop people's capabilities. In essence, 'responsabilisation' means shifting more operational responsibility to those people who work on the factory floor. It also involves learning new skills: how to work effectively in teams, how to structure projects, how to manage conflicts, and how to communicate in non-confrontational ways. Rather than issue direct instructions, team leaders act as coaches or, if any conflict arises, as referees. Workers in the teams allocate responsibility between themselves for tasks such as production scheduling, safety procedures, quality control, and so on. After the company introduced the idea of greater worker autonomy in one of its factories, staff there said that they felt happier and more productive as a result. The company felt it was such a success that it extended the practice to six factories in Europe and North America.

- understand the practicalities of the improvement process and be able to communicate its principles and techniques to the rest of the organisation;
- be able to participate in the total problem-solving process to improve performance;
- formulate and maintain a clear idea of the operation's improvement philosophy.

Should the improvement process be formally supervised?

Some improvement processes fail because they develop an unwieldy 'bureaucracy' to run them. But any process needs to be managed, so all improvement processes will need some kind of group to design, plan and control its efforts. However, a worthwhile goal for many improvement processes is to make them 'self-governing' over time. In fact, there are significant advantages in terms of people's commitment in giving them responsibility for managing the improvement process. However, even when improvement is driven primarily by self-managing improvement groups, there is a need for some sort of 'repository of knowledge' to ensure that the learning and experience accumulated from the improvement process is not lost.

To what extent should improvement be group based?

No one can really know a process quite like the people who operate it. They have access to the informal as well as the formal information networks that contain the way processes really work. But working alone, individuals cannot pool their experience or learn from one another, so improvement processes are almost always based on teams. The issue is how these teams should be formulated, which will depend on the circumstances of the operation, its context and its objectives. For example, *quality circles*, much used in Japan, have encountered mixed success in the West. A very different type of team is the '*task force*', or what some US companies call a 'tiger team'. Compared to quality circles, this type of group is far more management directed and focused. Most improvement teams are between these two extremes.

How should success be recognised?

If improvement is so important it should be recognised, with success, effort and initiative being formally rewarded. The paradox is that, if improvement is to become part of everyday operational life, why should improvement effort be especially rewarded? One compromise is to devise a recognition and rewards system that responds to improvement initiatives early in the improvement process, but then merges into the operation's normal reward procedures. In this way, people are rewarded not just for the efficient and effective running of their processes on an ongoing basis, but also for improving their processes. Improvement will then become an everyday responsibility of all people in the operation.

How much training is required?

Training has two purposes in the development of improvement processes. The first is to provide the necessary skills that will allow staff to solve process problems and implement improvements. The second is to provide an understanding of the appropriate interpersonal, group and organisational skills that are needed to 'lubricate' the improvement process. This second objective is more difficult than the first. Training and improvement techniques may take up significant time and effort, but none of this knowledge will be of much use if the organisational context for improvement militates against the techniques being used effectively. Although the nature of appropriate organisational development is beyond the scope of this text, it is worth noting that both technique-based skills and organisational skills are enhanced if staff have a basic understanding of the core ideas and principles of operations and process management.

Critical commentary

- Many of the issues covered in this chapter are controversial, for different reasons. Some criticism concerns the effectiveness of improvement methods. For example, it can be argued that there is a fundamental flaw in the concept of benchmarking. Operations that rely on others to stimulate their creativity, especially those that are in search of 'best practice', are always limiting themselves to currently accepted methods of operating or currently accepted limits to performance. 'Best practice' is not 'best' in the sense that it cannot be bettered, it is only 'best' in the sense that it is the best one can currently find. And accepting what is currently defined as 'best' may prevent operations from ever making the radical breakthrough or improvement that takes the concept of 'best' to a new and fundamentally improved level. Furthermore, because one operation has a set of successful practices in the way it manages its process does not mean that adopting those same practices in another context will prove equally successful. It is possible that subtle differences in the resources within a process (such as staff skills or technical capabilities) or the strategic context of an operation (for example, the relative priorities of performance objectives) will be sufficiently different to make the adoption of seemingly successful practices inappropriate.

- Other approaches are seen by some as too radical and too insensitive. For example, business process re-engineering has aroused considerable controversy. Most of its critics are academics, but some practical objections to BPR have also been raised, such as the fear that BPR looks only at work activities rather than at the people who perform the work. Because of this, people become 'cogs in a machine'. Also, some see BPR as being too imprecise because its proponents cannot agree as to whether it has to be radical or whether it can be implemented gradually, or exactly what a process is, or whether it has to be top-down or bottom-up, or whether it has to be supported by information technology or not.

Perhaps most seriously, BPR is viewed as merely an excuse for getting rid of staff. Companies that wish to 'downsize' (that is, reduce numbers of staff within an operation) are using BPR as an excuse. This puts the short-term interests of the shareholders of the company above either their longer-term interests or the interests of the company's employees. Moreover, a combination of radical redesign together with downsizing can mean that the essential core of experience is lost from the operation. This leaves it vulnerable to any market turbulence, since it no longer has the knowledge and experience of how to cope with unexpected changes.

- Even the more gentle approach of continuous improvement is not universally welcomed. Notwithstanding its implications of empowerment and liberal attitude towards shop-floor staff, it is regarded by some worker representatives as merely a further example of management exploiting workers. Its critics have defined relatively established ideas such as TQM as 'management by stress'. Or, even more radically, 'TQM is like putting a vacuum cleaner next to a worker's brain and sucking out ideas. They don't want to rent your knowledge any more, they want to own it – in the end that makes you totally replaceable.'

SUMMARY CHECKLIST

- ☐ Is the importance of performance improvement fully recognised within the operation?
- ☐ Do all operations and process managers see performance improvement as an integral part of their job?
- ☐ Is the gap between current and desired performance clearly articulated in all areas?
- ☐ Is the current performance measurement system seen as forming a basis for improvement?
- ☐ Does performance measurement focus on factors that reflect the operation's strategic objectives?
- ☐ Do performance measures allow likely problem areas to be diagnosed?
- ☐ Is some kind of balanced scorecard approach used that includes financial, internal, customer and learning perspectives?
- ☐ Is target performance set using an appropriate balance between historical, strategic, external and absolute performance targets?
- ☐ Are both performance and process methods benchmarked against similar operations and/or processes externally?
- ☐ Is benchmarking done on a regular basis and seen as an important contribution to improvement?
- ☐ Is some formal method of comparing actual and desired performance (such as the importance–performance matrix) used?
- ☐ To what extent does the operation have a predisposition towards breakthrough or continuous improvement?
- ☐ Have breakthrough improvement approaches such as business process re-engineering been evaluated?
- ☐ Are continuous improvement methods and problem-solving cycles used within the operation?
- ☐ If they are, has continuous improvement become a part of everyone's job?
- ☐ Has the Six Sigma approach to improvement been evaluated?
- ☐ Are the more common improvement techniques used to facilitate improvement within the operation?
- ☐ Does the operation show any signs of becoming a fashion victim of the latest improvement approach?
- ☐ Does the operation have a well-thought-through approach to managing improvement?

Case study

Ferndale Sands Conference Centre[16]

Mario Romano, the owner and General Manager of Ferndale Sands Conference Centre, had just seen an article in *The Conference Centre Journal*, and he was furious. The excellent reputation that he had worked so hard to build up over the last ten years was being threatened by one unreasonable customer and a piece of sloppy, sensationalist journalism. '*It really is unfair. Why do they let one mistake dictate the whole story? I'll tell you why, it's because they are more interested in a damning headline than they are in representing the truth*.'

Ferndale Sands Conference Centre is a conference venue of 52 rooms in the state of Victoria, Australia, about 20 km outside Melbourne. Established by Mario's father, initially as a hotel, it was relaunched as a conference centre four years ago and Mario was broadly pleased with the business he had attracted so far. The centre had managed to establish a presence in the fast-growing and profitable conference market. Specifically, it aimed at the 'executive retreat', rather than the 'company meeting' market. These events could be anything from one day's through to two weeks' duration. He had also negotiated deals with three higher education institutions to accommodate their executive education programmes. With its traditional 'Victorian' architecture, tranquil setting and excellent kitchen, Ferndale Sands offered a '. . . *supremely comfortable setting in which to work on those important decisions that will shape the future of your organisation*'. (Ferndale Sands brochure*)*

What had infuriated Marco was an article in the journal that had claimed to uncover administrative complacency and inefficiency at some of the state's conference centres (see the extract in Figure 12.16). Yet in the same edition, another piece had, generally, given a good rating to Ferndale Sands. This article had compared four conference centre facilities in and around Melbourne, and although the editorial comment had been neutral, the details included in the survey had quite clearly shown Ferndale Sands in a favourable light. Table 12.2 shows the summary of the four conference centres reviewed.

Both Mario and his Front-of-House Manager Robyn Wells disputed the article's rating of their administrative capabilities. However, they also were aware that administrative support was seen as being fairly important when they surveyed their clients (see Table 12.3 for Ferndale Sands' latest survey results).

Mario was determined to do something about the negative publicity. He called a meeting between himself, Nick Godfrey, who was in charge of catering and recreational facilities, and Robyn Wells, the Front-of House Manager, who also was in charge of all client relations.

Nick – '*OK Mario, I know you're not pleased, but I think you are in danger of overreacting. The best way to respond is just to ignore it. It's the survey that will be saved by potential clients, not a minor article at the front of the journal. And it's the survey that reflects what we really are*.'

Mario – '*Yeah, but even that gets it wrong. It shows administration and support as uneven. What do they mean uneven?*'

Robin – '*I don't know, I guess they must have talked to a couple of clients with some kind of grudge. But look, two things always come out as the most important things for our customers: quality of service and the flexibility to accommodate their needs for different configuration of rooms. We're great at service quality. We're always getting extravagant praise; it's a real winner for us. I've got files full of compliments. It's room flexibility that's our problem. Most clients accept that you can't mess around with a historic building like this, but that doesn't get round the fact that we can't reconfigure our rooms like you could in a modern hotel with sliding room partitions*.'

Mario – '*Well maybe that's something we could minimise by making it clear to clients what we can and can't do when they make a reservation*.'

Robyn – '*True, and we do that when demand is very high. But you can't ask us to turn away business by stressing what we can't do during quiet periods*.'

Shambles behind the grandeur?

Alison Peraway

Even the grandest of Victorian conference centres can fall from grace it seems. Recent complaints that Ferndale Sands may look like a Governor's palace, but can't get the basics right, were supported in hard hitting comments from the State's leading reservations agency. *"Ferndale Sands may not be the only venue to get complacent", said Charles.*

Figure 12.16 Extract from *The Conference Centre Journal*

Table 12.2 Extract from *The Conference Centre Journal*'s survey of conference venues in and around Melbourne

	Ferndale Sands Conference Centre	*Collins International*	*The Yarrold Conference Centre*	*St Kildan Conference Centre*
Price ($$$$$ = very expensive, $$$ = average, $ = budget)	$$$$	$$$$$	$$$	$
Size of menu	Extensive	Standard	Standard	Limited
Décor	Traditional, luxurious and tasteful	Modern, very luxurious and stylish	Modern but basic	Needs renovation
Style and quality of food	Modern, best in the state	International modern, slightly standardised	Varies, undistinguished but acceptable	Varies, very basic
Quality of service	Excellent, friendly and relaxed	Good	Limited	Enthusiastic but limited
Administration and support	Variable, some problems recently	Good	Good	OK
Flexibility of accommodation	Poor	Very good	Acceptable	OK
Off-peak price discounts	None offered	Some in summer, none in winter	Willing to negotiate	Willing to negotiate
Equipment	Normal range	Normal range	Normal range	Requires notice for 'anything unusual'
Recreational facilities	Full range: gym, tennis, golf, swimming pool, etc.	Gym, swimming pool	Gym only	Gym only
Ease of access	Good, will pick up from airport and city	City centre, no airport shuttle to hotel, but bus service, taxis, etc.	Close to city centre	10 km from city centre

Table 12.3 Percentage of Ferndale Sands clients reporting factor as important or very important

	Percentage of customers reporting factor as important or very important
Price	72
Size of menu	16
Décor	55
Style and quality of food	58
Quality of service	89
Administration and support	56
Flexibility of accommodation	85
Off-peak price discounts	16
Equipment	72
Recreational facilities	21
Ease of access	73

Mario – '*Well maybe we should. But that's not my main concern right now. What worries me are the things that always show up as mid-range factors in our customer surveys. We tend to forget about these. They may not be the most important things in the clients' eyes but they're not unimportant either. I'm talking about things like the quality of our food and the décor of the rooms, and also Robyn, the administrative support we offer. If we get these things wrong it can almost cause us as many problems as the really important things. That's why I'm upset about the poor administrative score we get in the journal. We score five for décor, and really good for food, but poor for administration.*'

Robyn – '*But as we said, that's just unfortunate. I still have every confidence in my administrative staff.*'

Nick – '*Before we get into that again, can I raise the question of our recreation facilities? It's one of our best assets yet it never rates as important with clients. It's the same with the choice we offer on our menus. Both these things are expensive to provide, and yet we don't seem to get the benefit. Why don't we make a real effort to* *really promote both of these things? You know, really convince the clients that great facilities and a wide choice on the menu are things that are worth paying a little more for.'*

Mario – '*It's not our pricing that's the problem. Although it's a fairly important issue with most clients, we can command relatively high prices. It's our costs that worry me more. Our general running costs are higher than they should be. Talking to the guys over at Parramatta Pacific in New South Wales; they are very similar to us, but their costs are a good 10 per cent less than ours.*'

Robyn – '*So, what is more important, raising our revenue or cutting our costs?*'

Mario – '*They are both equally important of course. The point is, what do we do about attracting more business and keeping our costs down?*'

Robyn – '*OK, we've got to do something, but remember we've also got the centre to run. Our busy period is just coming up and I don't want everyone distracted by lots of little improvement initiatives*.'

Nick – '*Absolutely. We have to limit ourselves to one or two actions that will have a noticeable impact.*'

Mario – '*I think you're probably right. But I would also add a further comment. And that is that if what we choose to do requires investment, then it must be guaranteed to have an impact. I need to go now, but why don't you two draw up a list of things that we could do. I'll review them later. OK, thanks everyone*.'

Robyn's and Nick's suggestions

In fact, Robyn and Nick decided to draw up individual lists of potential improvement initiatives. They also decided that, to begin with, only two of these improvement initiatives should be chosen.

Robyn's suggestions:

1. Increase prices – '*Why not? Although demand is variable, the general trend is rising as the conference market expands. At this top end of the market I don't think we are that price-sensitive. It would also bring in the revenue that we need to make further reinvestments to the centre*.'
2. Reduce menu choice – '*This really is a left-over from what menus used to be like. It goes back to the time when "more" was considered "better" instead of just "more". It is also very expensive to maintain that range of food while still maintaining quality*.'
3. Close the golf course – '*The golf course is probably the most expensive facility we have outside the house. It isn't rated by customers, so why do we keep it on?*'
4. Renovate outbuilding to provide flexible conference rooms – '*We can't easily change the inside of the house, but we do have outbuildings that could be converted into conference suites. They could be equipped with moveable partitions that would enable the space to be configured however our clients wanted it. OK, it would be expensive, but in the long term it's necessary*.'

Nick's suggestions:

1. Promote food and facilities more effectively – '*We have a great reputation for food and for having marvellous facilities. Ferndale Sands is just a beautiful place, yet we're not exploiting it fully. A campaign from a good public relations company could really establish us as the premier conference centre, not just in the state, but in the whole of Australia*.'
2. Cut in-house staff numbers and replace part-time staff with a smaller number of full-time staff – '*Having so many part-time staff is expensive. We pay them the same hourly rate as full-time staff yet there are all the extra personnel costs. Also, I think we are overstaffed in the house. Staff costs are a major part of our expenditure. It's the obvious area to look for cuts*.'
3. Invest in more equipment, both relaxation equipment and presentation equipment – '*We have great sporting facilities but they could be better. If we are going to exploit them more it may not be a bad thing to invest even more heavily in them. Also, we could make sure that we are ahead of the game in terms of the very latest audio-visual equipment. Both these things would help us to promote ourselves as the premier conference centre in Australia*.'

Klaus Vedfelt/DigitalVision/Getty Images

Questions

1. What factors would you use to judge the operations performance of Ferndale Sands?
2. What improvement priority would you give to each of these performance measures?
3. Which two suggestions put forward by Robyn and Nick would you recommend?

Applying the principles

Some of these exercises can be answered by reading the chapter. Others will require some general knowledge of business activity and some might require an element of investigation. **Model answers for the first two questions are available on the student companion website**. *For model answers to the other questions in this section, please ask your tutor/lecturer.*

1. '*Everything we do can be broken down into a process*', said Lucile, COO of an outsourcing business for the 'back-office' functions of a range of companies. '*It may be more straightforward in a manufacturing business, but the concept of process improvement is just as powerful in service operations. Using this approach our team of Black Belts has achieved 30 per cent productivity improvements in six months. I think Six Sigma is powerful because it is the process of comparing process outputs against customer requirements. To get processes operating at less than 3.4 defects per million opportunities means that you must strive to get closer to perfection and it is the customer that defines the goal. Measuring defects per opportunity means that you can actually compare the process of, say, a human resources process with a billing and collection process*.'

a) What are the benefits of being able to compare the number of defects in a human resources process with those of collection or billing?

b) Why is achieving defects of less than 3.4 per million opportunities seen as important by Lucile?

c) What do you think are the benefits and problems of training Black Belts and taking them off their present job to run the improvement projects, rather than the project being run by a member of the team that has responsibility for actually operating the process?

2. Develop cause–effect diagrams for the following types of problem:

a) Staff waiting too long for their calls to be answered by their IT helpdesk.

b) Poor food in the company restaurant.

c) Poor lecturing from teaching staff at a university.

d) Customer complaints that the free plastic toy in their breakfast cereal packet is missing.

e) Staff having to wait excessively long periods to gain access to the coffee machine.

3. For over 10 years a hotel group had been developing self-managed improvement groups within its hotels. At one hotel reception desk, staff were concerned about the amount of time the reception desk was left unattended. To investigate this, the staff began keeping track of the reasons they were spending time away from the desk and how long each absence kept them away. Everyone knew that reception desk staff often had to leave their post to help or give service to a guest. However, no one could agree what was the main cause of absence. Collecting the information was itself not easy because the staff had to keep records without affecting customer service. After three months, the data shown in Table 12.4 was presented to the staff in the form of a Pareto diagram. It came as a surprise to reception staff and hotel management that making photocopies for guests was the main reason for absence. Fortunately, this was easily remedied by moving the photocopier to a room adjacent to the reception area, enabling staff to keep a check on the reception desk while they were making copies.

a) Do you think it was wise to spend so much time on examining this particular issue? Isn't it a trivial issue?

b) Should this information be used to reflect improvement priorities? In other words, was the group correct to give priority to avoiding absence through photocopying, and should its next priority be to look at the time spent checking files in the back office?

4. A transport services company provides a whole range of services to railway operators. Its reputation for quality is a valuable asset in its increasingly competitive market. '*We are continually looking for innovation in the way we deliver our services, because the continuous improvement of our processes is the only way to make our company more efficient*', said

Table 12.4 Reasons for staff time away from the reception desk

Reason for being away from reception desk	*Total number of minutes away*
Checking files in back office	150
Providing glasses for night drinks	120
Providing extra key cards	90
Providing medication	20
Providing extra stationery	70
Providing misc. items to rooms	65
Providing night drinks	40
Making photocopies	300
Carrying messages to meeting rooms	125
Locking and unlocking meeting rooms	80
Providing extra bed linen	100

the company's CEO. '*We use a defined set of criteria to identify critical processes, each of which is allocated a "process owner" by our quality steering committee. This is helped by the company's "process excellence index" (PEI), which is an indicator of the way a process performs, particularly how it is designed, controlled and improved. The PEI score, which is expressed on a scale of 1–100, is calculated by the process owner and registered with the quality department. With this one figure we can measure the cost, reliability and quality of each process so that we can compare performance. If you don't measure, you can't improve. And if you don't measure in the correct way, how can you know where you are? Employee recognition is also important. Our suggestion scheme is designed to encourage staff to submit ideas that are evaluated and rated. No individual suggestion is finally evaluated until it has been fully implemented. Where a team of employees puts ideas forward, the score is divided between them, either equally or according to the wishes of the team itself. These employee policies are supported by the company's training schemes, many of which are designed to ensure all employees are customer-focused.*'

a) What seem to be the key elements in this company's approach to improvement?

b) Do you think this approach is appropriate for all operations?

5. a) As a group, identify a 'high-visibility' operation that you all are familiar with. This could be a type of quick-service restaurant, clothing store, public transport system, library, etc.

b) Once you have identified the broad class of operation, visit a number of them and use your experience as customers to identify:

i) the main performance factors that are of importance to you as customers; and

ii) how each operation rates against the others in terms of their performance on these same factors.

c) Draw an importance–performance matrix for one of the operations that indicates the priority it should be giving to improving its performance.

d) Discuss the ways in which such an operation might improve its performance and try to discuss your findings with the staff of that operation.

Notes on chapter

1 Carroll, L. (1871) *Through the Looking Glass*, Penguin Classics, 2003.

2 Lebas, M.J. (1995) 'Performance measurement and performance management', *International Journal of Production Economics*, 41 (1–3), pp. 23–35.

3 Kaplan, R.S. and Norton, D. (1996) *The Balanced Scorecard*, Harvard Business School Press.

4 The information on which this example is based is taken from: Economist (2016) 'The great escape: What other makers can learn from the revival of Triumph motorcycles', *Economist* print edition, 23 January.

5 Shenkar, O. (2010) *Copycats: How smart companies use imitation to gain a strategic edge*, Harvard Business School Press.

6 Shenkar, O. (2012) 'Just imitate it! A copycat path to strategic agility', *Ivey Business Journal*, May/June.

7 The information on which this example is based is taken from: Press release (2020) 'Mercedes F1 help develop coronavirus breathing device for health service use', Formula1.com, 30 March; West, K. (2011) 'Formula One trains van drivers', *The Times*, 1 May; Williams F1 Team (2010) 'Williams in collaboration with Sainsbury's', f1network.net, 12 November, http://www.f1network.net/main/s107/st164086.htm [accessed 28 September 2020].

8 Ferdows, K. and de Meyer, A. (1990) 'Lasting improvements in manufacturing performance: In search of a new theory', *Journal of Operations Management*, 9 (2), pp. 168–184. However, research for this model is mixed. For example, Patricia Nemetz questions the validity of the mode, finding more support for the idea that the sequence of improvement is generally dictated by technological (operations resource) or market (requirements) pressures: Nemetz, P. (2002) 'A longitudinal study of strategic choice, multiple advantage, cumulative model and order winner/qualifier view of manufacturing strategy', *Journal of Business and Management*, January.

9 Hammer, M. and Champy, J. (1993) *Re-engineering the Corporation: A manifesto for business revolution*, Nicholas Brealey Publishing.

10 Hammer, M. (1990) 'Re-engineering work: Don't automate, obliterate', *Harvard Business Review*, 68 (4).

11 The information on which this example is based is taken from: Onetto, M. (2014) 'When Toyota met e-commerce: Lean at Amazon', *McKinsey Quarterly*, 21, February.

12 Pande, P.S., Neuman, R.P. and Cavanagh, R. (2013) *The Six Sigma Way: How to maximize the impact of your change and improvement efforts*, 2nd edition, McGraw-Hill Education.

13 The information on which this example is based is taken from: Harvin, H. (2020) 'Six Sigma training & implementation at Wipro', henryharvin.com blog, 2 March, https://www.henryharvin.com/blog/six-sigma-training-implementation-at-wipro/ [accessed 28 September 2020]; Sharma, M., Pandla, K. and Gupta, P. (2014) 'A case study on Six Sigma at Wipro Technologies: Thrust on quality', Working Paper, The Jaipuria Institute of Management; Wipro website, www.wipro.com [accessed 28 September 2020].

14 The information on which this example is based is taken from: Brandall, B. (2018) 'The Checklist Manifesto review', process.st blog, 1 August, https://www.process.st/

the-checklist-manifesto-review/; Gawande, A. (2010) *The Checklist Manifesto: How to get things right*, Profile Books.

15 The information on which this example is based is taken from: Hill, A. (2017) 'Power to the workers: Michelin's great experiment', *Financial Times*, 11 May; Hill, A. and Stothard, M. (2017) 'Michelin chief Jean-Dominique Senard devolves power to workers', *Financial Times*, 14 May; '2016 Michelin Annual Report' (2017).

16 This case is based on several residential conference centres but does not reflect the concerns of any particular one.

Taking it further

Ahlström, J. ***(2014) How to Succeed with Continuous Improvement: A primer for becoming the best in the world, McGraw-Hill Education.*** *Based on a real-life case study.*

Boutros, T. and Cardella, J. ***(2016) The Basics of Process Improvement, Routledge.*** *Deals with 'the basics', but insightful nonetheless.*

George, M., Rowlands, D. and Kastle, B. ***(2003) What Is Lean Six Sigma?, McGraw-Hill Education.*** *Very much a quick introduction to what Lean Six Sigma is and how to use it.*

Goldratt, E.Y. and Cox, J. ***(2004) The Goal: A process of ongoing improvement, 3rd edition, Gower Publishing Ltd.*** *Updated version of a classic.*

Harrington, H.J. ***(2011) Streamlined Process Improvement, McGraw-Hill Education.*** *Individualistic, but still manages to convey the essence of the main improvement approaches.*

Page, S. ***(2015) The Power of Business Process Improvement: 10 simple steps to increase effectiveness, efficiency, and adaptability, 2nd edition, AMACON.*** *Very much a 'how to do it' practitioner handbook, but none the worse for that.*

Pande, P.S., Neuman, R.P. and Cavanagh, R. ***(2013) The Six Sigma Way: How to maximize the impact of your change and improvement efforts, 2nd edition, McGraw-Hill Education.*** *Obviously based on the Six Sigma principle, this is an unashamedly practical guide to the Six Sigma approach.*

Zu, X., Fredendall, L.D. and Douglas, T.J. ***(2008) 'The evolving theory of quality management: The role of Six Sigma', Journal of Operations Management, 26 (5), pp. 630–650.*** *As it says. . .*

13 Quality management

Introduction

No business, or enterprise of any kind, can afford to ignore quality. It can build a company's reputation, allowing it to preserve its margins, and it can save the operational and reputational costs of poor quality. That is why all businesses are concerned with quality, usually because they have come to understand that high quality can be a significant competitive advantage. But 'quality management' has come to mean more than avoiding errors. It is also seen as an approach to the way processes should be managed and, more significantly, improved generally. This is because quality management focuses on the very fundamental of operations and process management – the ability to produce and deliver the products and services that the market requires, both in the short and long terms. A grasp of quality management principles is the foundation of any improvement activity. Figure 13.1 shows the position of the ideas described in this chapter within the general model of operations management.

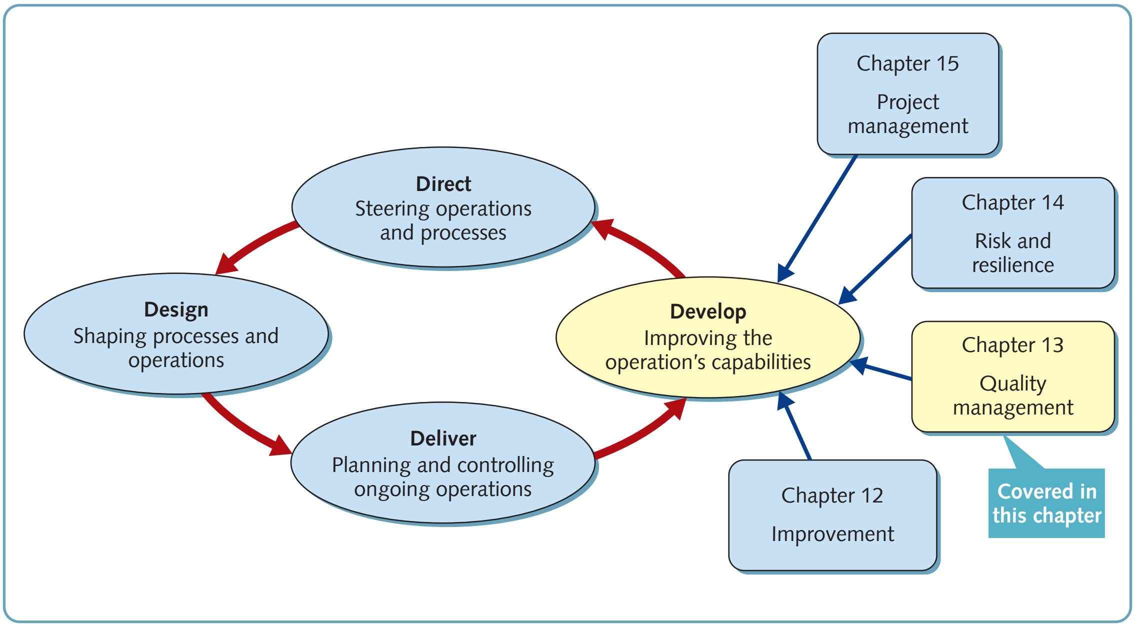
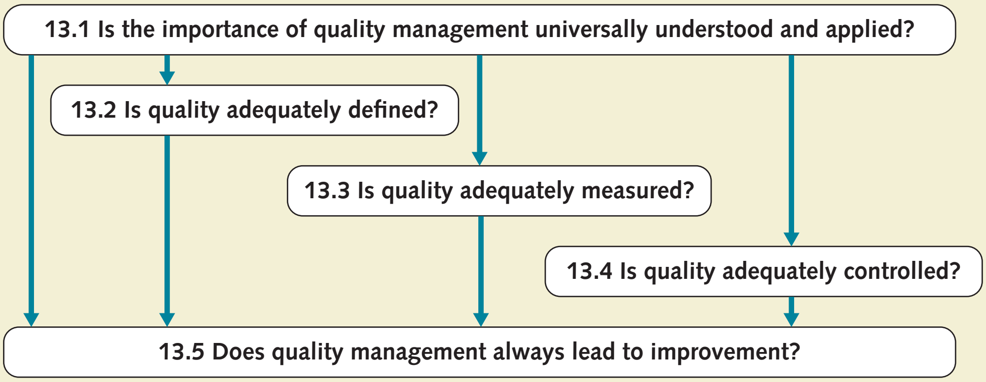


Figure 13.1 Quality management is the activity of ensuring consistent conformance to customers' expectations

EXECUTIVE SUMMARY



13.1 Is the importance of quality management universally understood and applied?

Quality is consistent conformance to customers' expectations. Managing quality means ensuring that an understanding of its importance, and the way in which it can be improved, is spread throughout the business. It is a subject that has undergone significant development over the last several decades. It has broadened its scope to include quality of service and quality of experience concepts. But arguably the most significant impact on how quality is managed has come from the total quality management (TQM) movement. Quality management is now seen as something that can be universally applied throughout a business and that also, by implication, is the responsibility of all managers in the business. In particular, it is seen as applying to all parts of the organisation. The internal customer concept can be used to establish the idea that it is important to deliver a high-quality service to internal customers (other processes in the business). Service-level agreements can be used to operationalise the internal customer concept. Just as important is the idea that quality also applies to every individual in the business.

13.2 Is quality adequately defined?

Quality needs to be understood from the customers' point of view because it is defined by the customers' perceptions and expectations. One way of doing this is to use a quality gap model. This starts from the fundamental potential gap between customers' expectations and perceptions, and deconstructs the various influences on those perceptions and expectations. Gaps between these factors can then be used to diagnose possible root causes of quality problems. A further development is to define the quality characteristics of products or services in terms of their functionality, appearance, reliability, durability, recovery and contact.

13.3 Is quality adequately measured?

Without measuring quality, it is difficult to control it. And the various attributes of quality can be measured either as a variable (measured on a continuously variable scale) or as an attribute (a binary 'acceptable or not acceptable' judgement). One approach to measuring quality is to express all quality-related issues in cost terms. Quality costs are usually categorised as prevention costs (incurred when trying to prevent errors), appraisal costs (associated with checking for errors), internal failure costs (errors that are corrected within the operation), and external failure costs (errors that are experienced by customers). Generally, it is held that increasing expenditure on prevention will bring a more-than-equivalent reduction in other quality-related costs.

13.4 Is quality adequately controlled?

Control means monitoring and responding to any deviations from acceptable levels of quality. One of the most common ways of doing this is through statistical process control (SPC). This technique not only attempts to reduce the variation in quality performance, so as to enhance process knowledge, but also is used to detect deviations outside the 'normal' range of quality variation.

13.5 Does quality management always lead to improvement?

Quality improvements often may not sustain because there is no set of systems and procedures to support and embed them within the operation's day-to-day routines. The best-known system for doing this is the ISO 9000 approach, adopted now throughout the world. Of the other systems, one of the most widely known is the EFQM Excellence Model. Once known only as the basis of the European Quality Award, it is now used extensively as a self-assessment tool that allows organisations to assess their own quality systems.

13.1 Diagnostic question: *Is the importance of quality management universally understood and applied?*

Before looking at the 'how' and 'why' of quality management, it is useful to understand what the term 'quality' can be taken to mean.

'Quality', 'quality of service' and 'quality of experience'

We have defined quality as the *consistent conformance to customers' expectations*. This is useful because it can be used to describe 'quality' for either physical products or intangible services, or any combination of the two. However, not all authorities or individual operations view 'quality' in this way, which can lead to some confusion. For example, the term 'quality of service' (QoS) is often used to describe how an operation serves its customers by combining what we call 'quality' with some or all of 'speed', dependability' and 'flexibility'. So, one could describe the 'quality' of a supermarket as including such factors as the quality of its goods, the cleanliness of its facilities and the courtesy of its staff. But, in assessing its QoS, the supermarket would probably want to include other factors, such as the speed of service, the predictability of opening times, stockouts, the range of goods available, and so on.

The limitation of QoS is that it may not capture the overall satisfaction with the service as perceived by the users. More useful, claim some providers of services, is to try to assess 'quality of experience' (QoE). Quality of experience is the overall acceptability of the service, as perceived subjectively by the end user. More formally, it can be described as the degree of delight or annoyance of the user of an application or service resulting from the fulfilment of his or her expectations with respect to the utility and/or enjoyment of the application or service in the light of the user's personality and current state.[1] QoE is clearly related to, but differs from, QoS in that it expresses, and focuses on, user satisfaction *both objectively and subjectively*.

QoS genrally includes the aspects of a service that are under the control of the operation creating the service, whereas QoE involves both the aspects governed by the operation and those that are a function of the individual customer, and the context in which the service is consumed. Figure 13.2 illustrates the relationship between these ideas, and Table 13.1 shows some typical factors that could be included in assessing the 'quality', 'quality of service' and 'quality of experience' of a supermarket operation and an online education service.

The QoE concept originated in telecommunications operations, information technology (IT) and consumer electronics. Yet its underlying principles have a far wider application. The QoE idea can be applied to any consumer-related business or service where the end user of a service or product could assess its quality subjectively and partly dependent on the context in which it is consumed. But the dependence of this judgement on the individual subjectivity of the user, and on the context of its consumption (which is beyond the influence of the operation) is both a strength and a weakness for the concept of QoE. An obvious strength is that it focuses operations on the richness of how their offerings are experienced by users. A practical weakness is that it is a difficult idea to operationalise. Subjective metrics of QoE are difficult to design, expensive and time-consuming.

OPERATIONS PRINCIPLE
Moving from 'quality', through 'quality of service', to 'quality of experience' shifts focus from 'the offering', through 'the provider', to 'the user'.

What is quality management?

There are many definitions of 'quality' – 'conformance to specification', being 'fit for purpose', 'achieving appropriate specification', and so on. The one we use here is '*quality is consistent conformance to customers' expectations*' because it includes both the idea of quality as

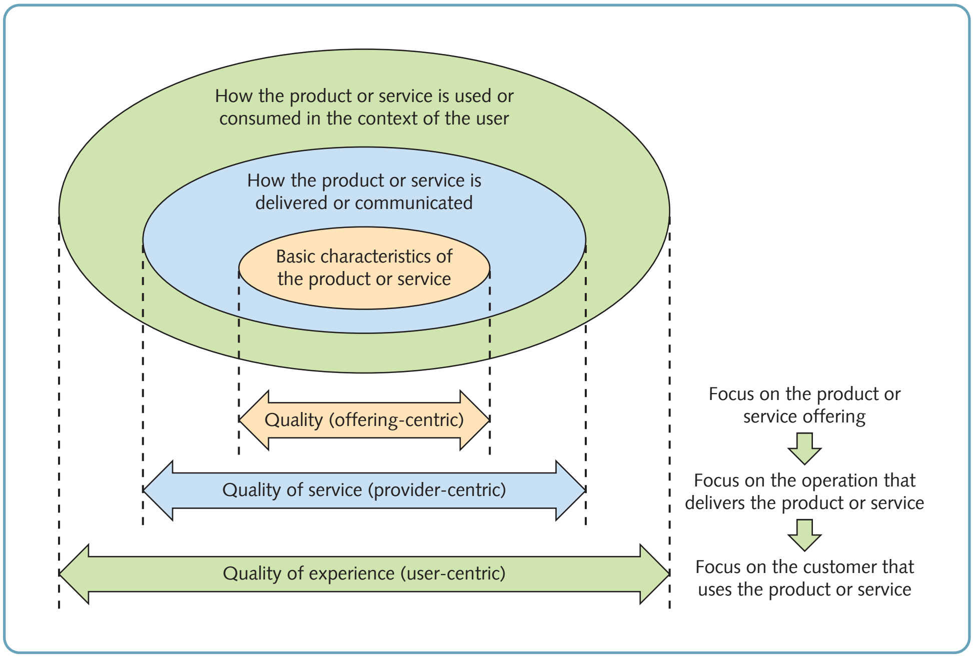


Figure 13.2 The relationship between 'quality', 'quality of service' and 'quality of experience'

Table 13.1 Examples of typical factors that could be included in assessing the 'quality', 'quality of service' and 'quality of experience' of a supermarket operation and an online education service

Operation	*Quality*	*Quality of service (QoS): 'quality' plus the following*	*Quality of experience (QoE): 'quality' plus 'quality of service' plus the following*
Supermarket	Quality of goods Cleanliness Staff courtesy	Speed of service Stockouts Predictability of service (hours of opening) Range of goods stocked	Open when I want to use the service? Perception of speed of service Is what I want available? Nature of other users of the service
Online education service	Quality/accuracy of the lesson content Quality of production values	Various technical measures of network performance (such as throughput, packet loss, delay and jitter)	Relevance of content to me How the content works on the device I am using (display fidelity, transport/stalling quality, etc.) How easy is it for me to navigate the content?

specification (what the product or service can do), and quality as *conformance* (there are no errors, so it always does what it is supposed to do). Not surprisingly, for such an important topic, it has a history. Arguably, the most significant of the approaches to quality management is 'total quality management' (TQM), which became popular with all types of business in the late 1970s and 1980s.

TQM can be viewed as a logical extension of the way in which quality-related practice has progressed. Originally quality was achieved by inspection – screening out defects before customers noticed them. Then the 'quality control' (QC) concept developed a more systematic approach to not only detecting, but also solving, quality problems. 'Quality assurance' (QA) widened the responsibility for quality to include functions other than direct operations, such as human resources, accounting and marketing. It also made increasing use of more sophisticated statistical quality techniques. TQM included much of what went before but developed its own distinctive themes, especially in its adoption of a more 'all-embracing' approach. Since the fashionable peak of TQM there has been some decline in its status, yet its ideas, many of which are included in this chapter, have become accepted quality practice.

Why is quality so important?

Put simply, quality is important because it has such a significant impact on both the revenues and costs of any business. Figure 13.3 illustrates the various ways in which quality improvements can affect other aspects of operations performance. Revenues can be increased by better sales and enhanced prices in the market. At the same time, costs can be brought down by improved efficiencies, productivity and the use of capital. So, a key task of the operations function must be to ensure that it provides quality goods and services, both to its internal and external customers.

Does quality apply to all parts of the organisation?

If quality management is to be effective, all processes must work properly together. This is because every process affects and in turn is affected by others. Called the *internal customer concept*, it is recognition that every part of an organisation is both an internal customer and, at the same time, an internal supplier for other parts of the organisation. This means that errors in the service provided within an organisation will eventually affect the product or service that reaches the external customer. So, one of the best ways of satisfying external customers is to satisfy internal customers. This means that each process has a responsibility to manage its own internal customer–supplier relationships by clearly defining its own and its customers' exact requirements. In fact, the exercise replicates what should be going on for the whole operation and its external customers.

> OPERATIONS PRINCIPLE
> ***An appreciation of, involvement in and commitment to quality should permeate the entire organisation.***

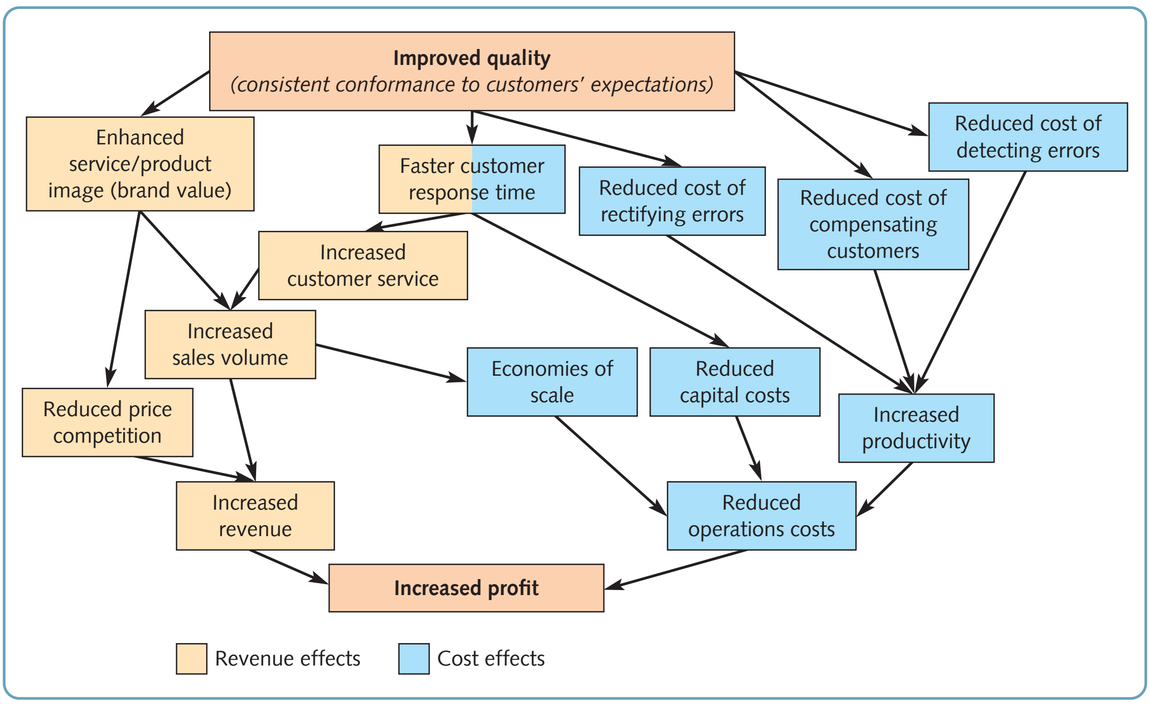


Figure 13.3 Higher quality has a beneficial effect on both revenues and costs

Service-level agreements and operating-/operational-level agreements

In Chapter 7, when discussing supply chain management, we described service-level agreements (SLAs). It is worth noting that, from a quality management perspective, they are a form of bringing a degree of formality to the internal customer concept. Perhaps more pertinent to quality is the operating-level (or operational, as they can sometimes be called, confusingly) agreement (OLA). Often confused with service-level agreements, an OLA is an agreement that defines how various groups within an operation propose to work together to deliver what has been specified in the service-level agreement. The difference between a service-level agreement and an operational-level agreement lies in where the boundary is drawn between the internal provider of a service and the internal customer. An SLA is the service that an internal customer should receive; an OLA is how the parts of the operation that produce that service will go about producing the service in practice.

Does every person in the organisation contribute to quality?

A total approach to quality should include every individual in the business. People are the source of both good and bad quality and it is everyone's personal responsibility to get quality right. This applies not only to those people who can affect quality directly and have the capability to make mistakes that are immediately obvious to customers – for example those who serve customers face to face or physically make products. It also applies to those who are less directly involved in producing products and services. The keyboard operator who miskeys data, or the product designer who fails to investigate thoroughly the conditions under which products will be used in practice, could also set in motion a chain of events that customers eventually see as poor quality. It follows that, if everyone has the ability to impair quality, they also have the ability to improve it – if only by 'not making mistakes'. But their contribution is expected to go beyond a commitment not to make mistakes; they are expected to bring something positive to the way they perform their jobs. Everyone is capable of improving the way in which they do their own jobs and practically everyone is capable of helping others in the organisation to improve theirs. Neglecting the potential inherent in all people is neglecting a powerful source of improvement.

Case example

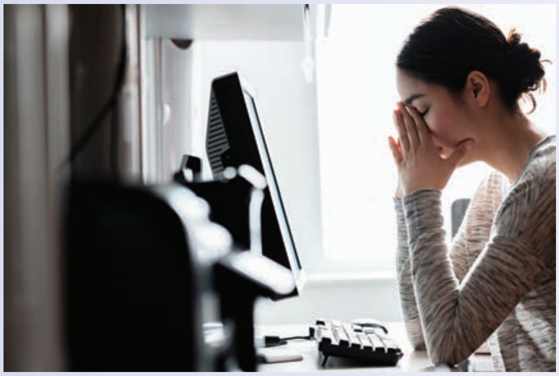
Keyboard errors – autofill and 'fat fingers'[2]

People make mistakes, especially at keyboards. Two types of mistake have proved both embarrassing and costly. The embarrassing one (which can also be costly) is the use of the autofill function on email and search applications. Fill in the first few letters, and the app does the rest for you. Unless it gets it wrong. Before the UK voted to leave the European Union, and when the topic was a politically and economically sensitive issue, the Bank of England's head of press mistakenly sent an email to the media revealing that officials were quietly researching the impact of Britain's exit from the EU. There were questions over whether the incident would cost him his job. Following the blunder, in a tightening of security regulations, Bank of England employees were prevented from using the autofill. Instead, staff were asked to write the full name of the recipient of their emails. It did not help productivity, but it reduced the possibility of an email getting to the wrong recipient.

The costly one (which can also be embarrassing) is what has become known as 'fat finger syndrome'. For example, feeling sleepy one day, a German bank worker briefly fell asleep on his keyboard when processing a €64 debit (withdrawal) from a pensioner's account, repeatedly pressing the number 2. The result was that the pensioner's account had €222 million withdrawn from it instead

▶

of the intended €64. Fortunately, the bank spotted the error before too much damage was done (and before the account holder noticed). More seriously, the supervisor who should have checked his junior colleague's work was sacked for failing to notice the blunder (unfairly, as a German labour tribunal later ruled). Fat finger trading mistakes are not uncommon. For example, the Swiss bank UBS mistakenly ordered 3 trillion yen (instead of 30 million yen) of bonds in a Japanese video games firm. In another example, a Japanese trader tried to sell one share of a recruitment company at 610,000 yen per share. But he accidentally sold 610,000 shares at one yen each, despite this being 41 times the number of shares available. Unlike the German example, the error was not noticed and the Tokyo Stock Exchange processed the order. It resulted in Mizuho Securities losing 27 billion yen. The head of the Tokyo Stock Exchange later resigned.



JGI/Tom Grill/Getty Images

13.2 Diagnostic question: *Is quality adequately defined?*

Quality is consistent conformance to customers' expectations. It needs to be understood from a customer's point of view because, to the customer, the quality of a particular product or service is whatever he or she perceives it to be. However, individual customers' expectations may be different. Past experiences, individual knowledge and history will all shape expectations. Perceptions are not absolute; exactly the same product or service may be perceived in different ways by different customers. Also, in some situations, customers may be unable to judge the 'technical' specification of the service or product. They may then use surrogate measures as a basis for their perception of quality. For example, after seeking financial advice from an adviser it might be difficult immediately to evaluate the technical quality of the advice, especially if no better solution presents itself. In reality a judgement of the quality of the advice may be based on perceptions of trustworthiness, relationship, the information that was provided or the way in which it was provided.

OPERATIONS PRINCIPLE
Perceived quality is governed by the magnitude and direction of the gap between customers' expectations and their perceptions of a product or service.

Closing the gap – alignment in quality

If the product or service experience was better than expected then the customer is satisfied and quality is perceived to be high. If the product or service was less than his or her expectations then quality is low and the customer may be dissatisfied. If the product or service matches expectations then the perceived quality of the product or service is seen to be acceptable. So:

When expectations > perceptions → perceived quality is poor

When expectations = perceptions → perceived quality is acceptable

When expectations < perceptions → perceived quality is good

Both customers' expectations and perceptions are influenced by a number of factors, some of which cannot be controlled by the operation and some of which can, at least to a certain extent. Figure 13.4 shows some of the factors that will influence the gap between expectations

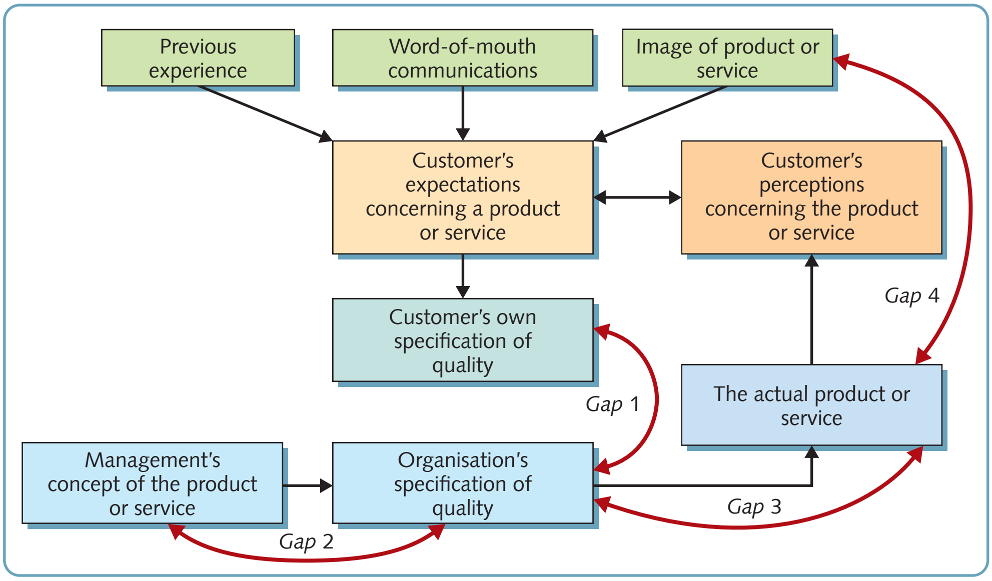


Figure 13.4 A perception–expectation gap model of quality

and perceptions and the potential gaps between some of these factors. This approach to defining quality is called a 'gap model of quality'. The model shown in Figure 13.4 is adapted from one developed by Zeithaml, Berry and Parasuraman,[3] primarily to understand how quality in service operations can be managed and to identify some of the problems in so doing. However, this approach is now used in all types of operation.

Diagnosing quality problems

Describing perceived quality in this way allows a diagnosis of quality problems. If the perceived quality gap is such that customers' perceptions of the product or service fail to match their expectations of it, then the reason (or reasons) must lie in other gaps elsewhere in the model. Four other gaps could explain a perceived quality gap between customers' perceptions and expectations.

Gap 1: The customer's specification–operation's specification gap

Perceived quality could be poor because there may be a mismatch between the organisation's own internal quality specification and the specification that is expected by the customer. For example, a car may be designed to need servicing every 10,000 kilometres, but the customer may expect 15,000-kilometre service intervals. An airline may have a policy of charging for drinks during the flight, whereas the customer's expectation may be that the drinks will be free.

Gap 2: The concept–specification gap

Perceived quality could be poor because there is a mismatch between the product or service concept and the way the organisation has specified the quality of the product or service internally. For example, the concept of a car might have been for an inexpensive, energy-efficient means of transportation, but the inclusion of a catalytic converter may have both added to its cost and made it less energy efficient.

Gap 3: The quality specification–actual quality gap

Perceived quality could be poor because there is a mismatch between the actual quality of the service or product provided by the operation and its internal quality specification. This may be the result, for example, of an inappropriate or unachievable specification, or of poorly trained or inexperienced personnel, or because effective control systems are not in place to ensure the provision of defined levels of quality. For example, if, despite an airline's policy of charging for drinks, some flight crews provide free drinks, they add unexpected costs to the airline and influence customers' expectations for the next flight, when they may be disappointed.

Gap 4: The actual quality–communicated image gap

Perceived quality could also be poor because there is a gap between the organisation's external communications or market image and the actual quality of the service or product delivered to the customer. This may be either the result of market positioning setting unachievable expectations in the minds of customers, or operations not providing the level of quality expected by the customer. The advertising campaign for an airline might show a cabin attendant offering to replace a customer's shirt on which food or drink has been spilt, but this service may not always be available.

Quality characteristics

Much of the 'quality' of a product or service will have been specified in its design. But not all the design details are useful in defining quality. Rather it is the *consequences* of the design that are perceived by customers. These consequences of the design are called *quality characteristics*. Table 13.2 shows a list of generally useful quality characteristics applied to a service (flight) and a product (car).

Table 13.2 Quality characteristics for a motor car and an air journey

Quality characteristics	*Car*	*Flight*
Functionality – how well the product or service does its job, including its performance and features	Speed, acceleration, fuel consumption, ride quality, road holding, etc.	Safety and duration of journey, onboard meals and drinks, car and hotel booking services
Appearance – the sensory characteristics of the product or service: its aesthetic appeal, look, feel, sound and smell	Aesthetics, shape, finish, door gaps, etc.	Décor and cleanliness of aircraft, look of waiting lounges and crew
Reliability – the consistency of the product's or service's performance over time, or the average time for which it performs within its tolerated band of performance	Mean time to failure	Keeping to the published flight times
Durability – the total useful life of the product or service assuming occasional repair or modification	Useful life (with repair)	Keeping up with trends in the industry
Recovery – the ease with which problems with the product or service can be rectified or resolved	Ease of repair	Resolution of service failures
Contact – the nature of the person-to-person contact that might take place; could include the courtesy, empathy, sensitivity and knowledge of contact staff	Knowledge and courtesy of sales and service staff	Knowledge, courtesy and sensitivity of airline staff

13.3 Diagnostic question: *Is quality adequately measured?*

Some quality characteristics are relatively easy to measure – for example, is the gap between a car door and pillar less than 5mm? Other more difficult-to-measure quality characteristics, such as 'appearance', need to be deconstructed into their constituent elements such as 'colour match', 'surface finish' and 'the number of visible scratches', all of which are capable of being measured in a relatively objective manner. They may even be quantifiable. However, deconstructing quality characteristics into their measurable sub-components can result in some loss of meaning. A quantified list of 'colour match', the 'smoothness' of the surface finish and the 'number of visible scratches' does not cover factors such as 'aesthetics' – a characteristic that is difficult to measure, but nonetheless important. Some quality characteristics cannot be measured at all. The 'courtesy' of airline staff, for example, has no objective quantified measure, yet airlines place a great deal of importance on the need to ensure courtesy in their staff. In cases like this, the operation will have to attempt to measure customer *perceptions* of the quality characteristic of courtesy.

Variables and attributes

The measures used to describe quality characteristics are of two types: variables and attributes. Variable measures are those that can be measured on a continuously variable scale (for example, length, diameter, weight or time). Attributes are those that are assessed by judgement and have two states (for example, right or wrong, works or does not work, looks OK or not OK). Table 13.3 categorises some of the measures that might be used for the quality characteristics of the car and the flight.

Table 13.3 Variable and attribute measures for quality characteristics

	Car		*Flight*	
Characteristic	*Variable*	*Attribute*	*Variable*	*Attribute*
Functionality	Acceleration and braking characteristics from test bed	Is the ridge quality satisfactory?	Number of journeys that actually arrived at the destination (i.e. didn't crash!)	Was the food acceptable?
Appearance	Number of blemishes visible on car	Is the colour to specification?	Number of seats not cleaned satisfactorily	Is the crew dressed smartly?
Reliability	Average time between faults	Is the reliability satisfactory?	Proportion of journeys that arrived on time	Were there any complaints?
Durability	Life of the car	Is the useful life as predicted?	Number of times service innovations lagged competitors'	Generally, is the airline updating its services in a satisfactory manner?
Recovery	Time from fault discovered to fault repaired	Is the serviceability of the car acceptable?	Proportion of service failures resolved satisfactorily	Do customers feel that staff deal satisfactorily with complaints?
Contact	Level of help provided by sales staff (1–5 scale)	Did customers feel well served (yes or no)?	The extent to which customers feel well treated by staff (1–5 scale)	Did customers feel that the staff were helpful (yes or no)?

Case example

Testing cars (close) to destruction[4]

Away from the public eye, at Millbrook Proving Ground, one of Europe's leading independent technology centres for the design, engineering, test and development of automotive and propulsion systems, they treat cars really badly. But all in a good cause. It is where auto manufacturers send their new vehicles to be tested, so that any glitches, from irritating rattles to more serious safety problems, can be exposed and corrected before the product reaches the market. The site, in Bedfordshire in the UK, is hidden away behind security fences and high embankments to discourage automobile paparazzi from taking photographs of new models as they are put through their paces. Auto manufacturers also test their new models out on public roads, usually with stick-on panels to disguise them, but for repeatable, carefully measured conditions, a facility like the Millbrook Proving Ground is needed. The site has been called 'an automotive time machine', where a gleaming new model drives in, and about 20 weeks later it drives out (if it can) having been exposed to the equivalent of 10 years of severe weather, and wear-and-tear comparable to being driven around 160,000 miles. During this time, it will have been driven on straight and twisty roads, up and down hills, slowly and very fast, through salt-water baths (to accelerate rusting) and along gravel roads that damage its paintwork. But that's not all. It will be roasted at high temperatures, frozen down to arctic conditions and drenched in water to expose any leaks. Also, it will be subjected to the infamous 'Belgian Pavé'. This is a mile-long track made from blocks of paving with rough sections and random depressions. The suspension takes such a beating that after five laps on the track vehicles need to be doused in a water trough to cool the dampers down. And during all this wrecking treatment, engineers are examining the vehicles periodically for signs of the wear or damage, to allow carmakers to fine-tune their designs or manufacturing processes in order to avoid the failures that would be expensive and reputationally damaging if they occurred after product launch.

High Level/Shutterstock

Measuring the 'costs of quality'

One approach to measuring aggregated quality is to express all quality-related issues in cost terms. This is the 'cost of quality' approach (usually taken to refer to both costs and benefits of quality). These costs of quality are usually categorised as *prevention costs*, *appraisal costs*, *internal failure costs* and *external failure costs*. Table 13.4 illustrates the types of factors that are included in these categories.

Understanding the relationship between quality costs

At one time it was assumed that failure costs reduce as the money spent on appraisal and prevention increases. There must be a point beyond which the cost of improving quality gets larger than the benefits that it brings. Therefore, there must be an optimum amount of quality effort to be applied in any situation that minimises the total costs of quality. Figure 13.5(a) sums up this idea.

More recently, the 'optimum-quality effort' approach has been challenged.[5] First, why should any operation accept the *inevitability* of errors? Some occupations seem to be able to accept a zero-defect standard (even if they do not always achieve it). No one accepts the inevitability of pilots crashing a certain proportion of their aircraft, or nurses dropping a certain number of babies. Second, failure costs are generally underestimated. They are usually taken to include the cost of 'reworking' defective products, 're-serving' customers, scrapping parts and materials, the

Table 13.4 The categories of quality cost

Category of quality-related cost	*Factors in these categories*
Prevention costs – those costs incurred in trying to prevent problems, failures and errors from occurring in the first place	• Identifying potential problems and putting the process right before poor quality occurs • Designing and improving the design of products and services and processes to reduce quality problems • Training and development of personnel in the best way to perform their jobs • Process control
Appraisal costs – those costs associated with controlling quality to check to see if problems or errors have occurred during and after the creation of the product or service	• The setting up of statistical acceptance sampling plans • The time and effort required to inspect inputs, processes and outputs • Obtaining processing inspection and test data • Investigating quality problems and providing quality reports • Conducting customer surveys and quality audits
Internal failure costs – failure costs that are associated with errors dealt with inside the operation	• The cost of scrapped parts and materials • Reworked parts and materials • The lost production time as a result of coping with errors • Lack of concentration due to time spent on troubleshooting rather than improvement
External failure costs – failure costs that are associated with errors being experienced by customers	• Loss of customer goodwill affecting future business • Aggrieved customers who may take up time • Litigation (or payments to avoid litigation) • Guarantee and warranty costs • The cost to the company of providing excessive capability (too much coffee in the pack, or too much information to a client)

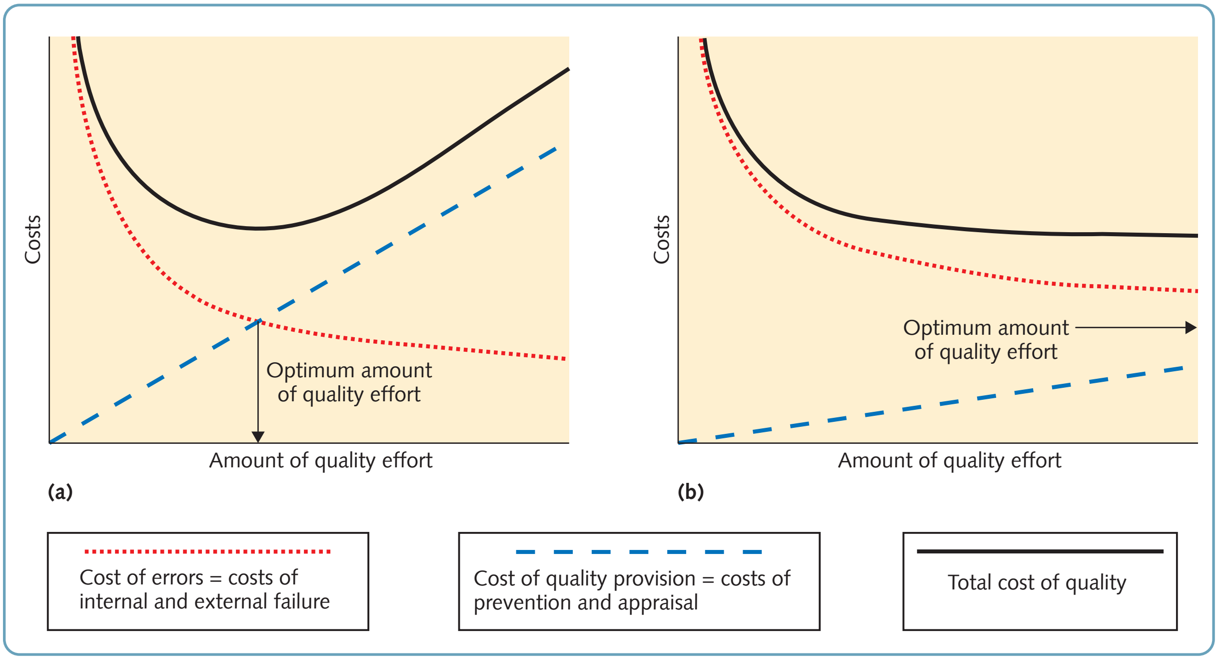


Figure 13.5 (a) The traditional 'cost of quality' model and (b) a more modern view

loss of goodwill, warranty costs, etc. These are important, but the real cost of poor quality should include all the management time wasted in organising rework and rectification and, more importantly, the loss of concentration and the erosion of confidence between processes within the operation. Third, it implies that prevention costs are inevitably high. But by stressing the importance of quality to every individual, preventing errors becomes an integral part of everyone's work. All have a responsibility for their own quality and all should be capable of 'doing things right first time'. This may incur some costs – training, automatic checks, anything that helps to prevent errors occurring in the first place – but not such a steeply inclined cost curve as in the 'optimum-quality' theory. Finally, the 'optimum-quality level' approach, by accepting compromise, does little to challenge operations managers and staff to find ways of improving quality.

Put these corrections into the optimum-quality effort calculation and the picture looks very different (see Fig. 13.5(b)). If there is an 'optimum', it is a lot further to the right, in the direction of putting more effort (but not necessarily cost) into quality.

The TQM-influenced quality cost model

TQM rejected the optimum-quality level concept. Rather, it concentrated on how to reduce all known and unknown failure costs. So, rather than placing most emphasis on appraisal (so that 'bad products and service don't get through to the customer'), it emphasised prevention (to stop errors happening in the first place). This has a significant positive effect on internal failure costs, followed by reductions in both external failure costs and, once confidence has been firmly established, in appraisal costs. Eventually, even prevention costs can be stepped down in absolute terms, though prevention remains a significant cost in relative terms. Figure 13.6 illustrates this idea, showing how total quality costs may rise initially as investment in some aspects of prevention is increased. This shifts the emphasis from a reactive approach to quality (waiting for errors to happen, then screening them out), to a more proactive 'getting it right first time' approach (doing something before errors happen).

OPERATIONS PRINCIPLE
Effective investment in preventing quality errors can significantly reduce appraisal and failure costs.

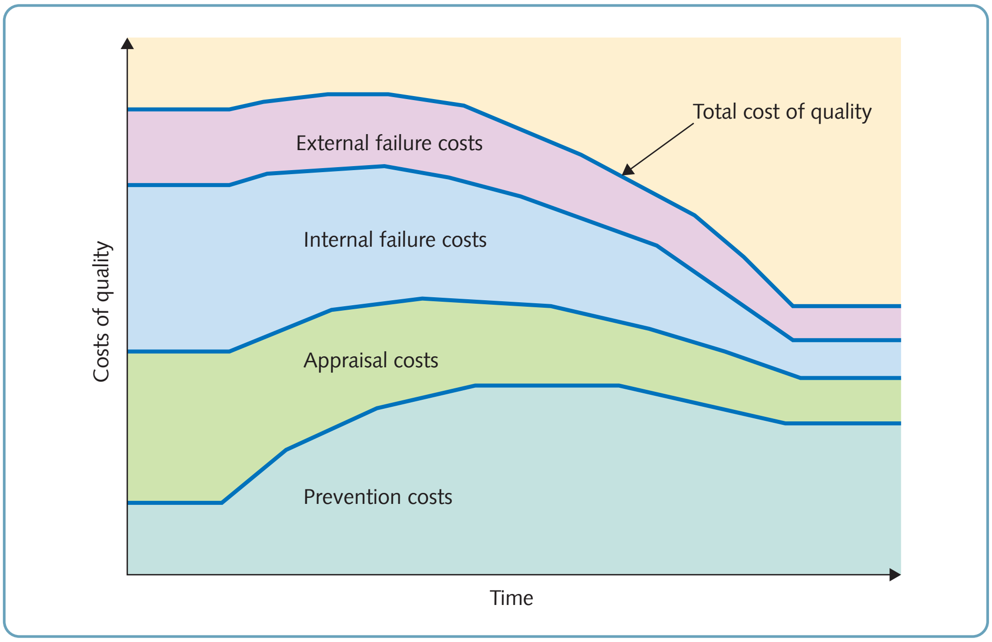


Figure 13.6 Increasing the effort spent on preventing errors occurring in the first place brings a more-than-equivalent reduction in other cost categories

13.4 Diagnostic question: *Is quality adequately controlled?*

After quality has been defined and measured, processes will need to check that their quality conforms to whatever quality standards are deemed appropriate. This does not necessarily mean checking everything; sampling may be more appropriate.

Check every product and service or take a sample?

There are several reasons why checking everything may not be sensible. For example, it could be dangerous to check everything – a blood test is just a small sample. Or checking of everything might destroy the product or interfere with the service – the life of every light bulb leaving the factory cannot be checked. Or checking everything may be too costly. Even 100 per cent checking will not always guarantee that all defects or problems will be identified. Inspections can miss mistakes, checks may be inherently difficult and information may be unreliable.

Sometimes, however, although expensive, 100 per cent inspection is necessary – for example, if a product is so critical that its failure would result in death or injury. Some aircraft parts, or some healthcare services are like this. In other cases, it may be that the economics of a 100 per cent inspection are such that the cost of doing it is relatively small. For example, some labels can be automatically scanned as they are produced, at virtually no extra cost. Yet, 100 per cent inspection can lead to another risk – that of classifying something as an error when it actually conforms to specification. This distinction is summarised in what are often referred to as type I and type II errors.

Type I and type II errors

Checking quality by sampling, although requiring less time than checking everything, does have its own problems. Take the example of someone waiting to cross a street. There are two main options: cross (take some action), or continue waiting (take no action). If there is a break in the traffic and the person crosses, or if that person continues to wait because the traffic is too dense, then a correct decision has been made (the action was appropriate for the circumstances). There are two types of incorrect decisions, or errors. One would be a decision to cross (take some action) when there is not an adequate break in the traffic, resulting in an accident – this is referred to as a type I error. Another would be a decision not to cross even though there was an adequate gap in the traffic – this is called a type II error. Type I errors are those that occur when a decision was made to do something and the situation did not warrant it. Type II errors are those that occur when nothing was done, yet a decision to do something should have been taken, as the situation did indeed warrant it. So, there are four outcomes, summarised in Table 13.5.

Table 13.5 **Type I and type II errors for a pedestrian crossing the road**

	Road conditions	
Decision	***Safe*** *(action was appropriate)*	***Unsafe*** *(action was not appropriate)*
Cross (take some action)	Correct decision	***Type I error***
Wait (take no action)	***Type II error***	Correct decision

Case example

Coin-counting calculations[6]

The rise in the use of contactless payment cards was always going to undermine the usefulness of physical money, a trend that the coronavirus pandemic accelerated. Even before that, the British Treasury had, for the first time in almost half a century, ordered the Royal Mint (who make coins) to stop producing any low-denomination coins. But what should people do with all their not-totally-unwanted coins? Most retail banks are reluctant to accept a large amount of coins unless they are counted and bagged in standard bags. However, Metro Bank, unlike any other bank in the UK, have free coin counters, called 'Magic Money Machines'. Yet, like all technology, coin-counting machines exhibit a certain amount of variation. An investigation showed that this variation can mean that the stated worth of the coins put into the machine can range from being perfectly accurate to the very last penny, through to as much as 19 per cent inaccurate. A journalist laboriously sorted and counted by hand £600-worth of coins – 14,500 coins in total. He then divided them into bags of coins that were worth exactly £100. He then visited Metro Bank branches in central and west London to see if the machines would count the coins as accurately as he had. In fact, the Magic Money Machines came out of the test reasonably well. Most were accurate, with a margin of error less than 1 per cent, which is pretty good. Even the machine that was 19 per cent out erred in the journalist's favour. In total, he ended up with a net profit of about £30 (the money was donated to Metro Bank's charity partner). Yet even small errors in such machines can add up in absolute terms when one considers that £22.5 million was processed by Metro Bank coin machines in the year prior to the investigation. But not all coin counters have proved as accurate as Metro Bank's. In the US, TD Bank had to abandon its 'Penny Arcade' coin-counting machines after widespread complaints that the devices were short-changing customers. An investigation had concluded that Penny Arcades in five locations inaccurately counted $300 packets of coins. And in no location did the counting error favour the customer.

Statistical process control (SPC)

The most common method of checking the quality of a sampled product or service so as to make inferences about all the output from a process is called statistical process control (SPC). SPC is concerned with sampling the process during the production of goods or the delivery of service. Based on this sample, decisions are made as to whether the process is 'in control' – that is, operating as it should be. If there seems to be a problem with the process, then it can be stopped (if possible and appropriate) and the problem identified and rectified. For example, an international airport may regularly ask a sample of customers if the cleanliness of its restaurants is satisfactory. If an unacceptable number of customers in one sample are found to be unhappy, airport managers may have to consider improving the procedures in place for cleaning tables.

Control charts

The value of SPC is not just to make checks of a single sample but to monitor the results of many samples over a period of time. It does this by using control charts. Control charts record some aspect of quality (or performance generally) over time to see if the process seems to be performing as it should (called *in control*), or not (called *out of control*). If the process does seem to be going out of control, then steps can be taken *before* there is a major problem.

Figure 13.7 shows a typical control chart. Charts something like these can be found in almost any operation. They could, for example, represent the percentage of customers in a sample of 1,000 who, each week, were dissatisfied with the service they received from two call centres. In chart (a), measured customer dissatisfaction has been steadily increasing over time. There is evidence of a clear (negative) trend that management may wish to investigate. In chart (b), although there is little evidence of any trend in average dissatisfaction, the variability in performance seems to be increasing. Again, the operation may want to investigate the causes.

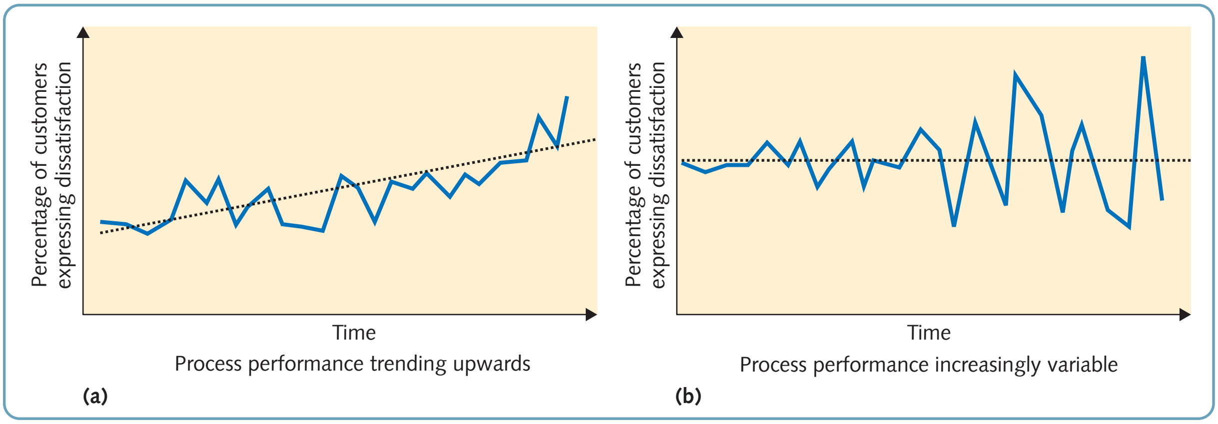


Figure 13.7 Control charting – any aspect of the performance of a process is measured over time and may show trends in average performance and/or changes in the variation of performance over time

Looking for *trends* is an important use of control charts. If the trend suggests the process is getting steadily worse, then it will be worth investigating the process. If the trend is steadily improving, it may still be worthy of investigation to try to identify what is happening that is making the process better. An even more important use of control charts, however, is to investigate the *variation* in performance.

Why variation is a bad thing

Although a trend such as that shown in Figure 13.7(a) clearly indicates deteriorating performance, the variation shown in Figure 13.7(b) can be just as serious. Variation is a problem because it masks any changes in process behaviour. Figure 13.8 shows the performance of two processes, both of which change their behaviour at the same time. The process on the left has such a wide natural variation that it is not immediately apparent that any change has taken place. Eventually it will become apparent, but it may take some time. By contrast, the performance of the process represented by the chart on the right has a far narrower band of variation, so the same change in average performance is more easily noticed. The narrower the variation of a process, the more obvious are any changes that might occur, and the easier it is to make a decision to intervene. SPC is discussed much further in the supplement to this chapter. It is also one of the core ideas in the Six Sigma improvement approach that was discussed in Chapter 12.

> OPERATIONS PRINCIPLE
> ***High levels of variation reduce the ability to detect changes in process performance.***

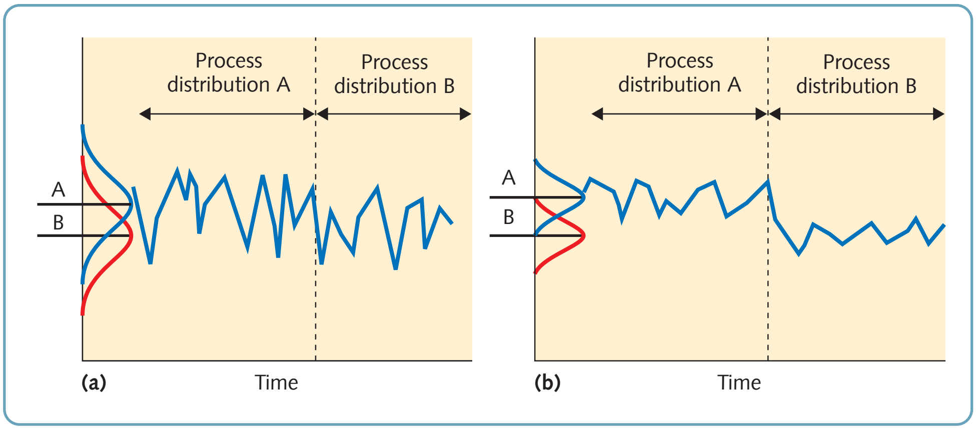


Figure 13.8 Low process variation allows changes in process performance to be more readily detected

Case example

What a giveaway[7]

Another negative effect of wide process variability is particularly evident in any process that fills any kind of container with weighed product. Even small reductions in the variability in filling levels can translate into major savings. This is because of what is known as 'giveaway' or 'over-fill', caused by the necessity to ensure that containers are not legally underweight. Although slightly different regulations may apply in various parts of the world, any process that produces products that have an 'e' after the stated weight on the container must produce products with an average weight greater than the declared weight on the container, with the average weight being determined by sampling. In addition, there are two other legal conditions. First, no more than 2.5 per cent of the sample can lie between an upper and lower control limit. Second, no product under a lower control limit may be produced at all. Therefore, operations managers often build in a margin of safety in order to overcome that variation, while allowing them to meet legal weight and measure conditions on minimum fill levels. As a result, containers on filling lines are routinely overfilled with more finished product than need be the case. This is why minimising fill variation can avoid 'giveaway'.

Process control, learning and knowledge

In recent years, the role of process control, and SPC in particular, has changed. Increasingly, it is seen not just as a convenient method of keeping processes in control, but also as an activity that is fundamental to the acquisition of competitive advantage. This is a remarkable shift in the status of SPC. Traditionally it was seen as one of the most *operational*, immediate and 'hands-on' operations management techniques. Yet it is now seen as contributing to an operation's *strategic* capabilities. This is how the logic of the argument goes:

1. SPC is based on the idea that process variability indicates whether a process is in control or not.
2. Processes are brought into *control* and improved by progressively reducing process variability. This involves eliminating the assignable causes of variation.
3. One cannot eliminate assignable causes of variation without gaining a better understanding of how the process operates. This involves *learning* about the process, where its nature is revealed at an increasingly detailed level.
4. This learning means that *process knowledge* is enhanced, which in turn means that operations managers are able to predict how the process will perform under different circumstances. It also means that the process has a greater capability to carry out its tasks at a higher level of performance.
5. This increased *process capability* is particularly difficult for competitors to copy. It cannot be bought 'off the shelf'. It comes only from time and effort being invested in controlling operations processes. Therefore, process capability leads to *strategic advantage*.

OPERATIONS PRINCIPLE
Statistical process control (SPC) gives the potential to enhance process knowledge.

13.5 Diagnostic question: *Does quality management always lead to improvement?*

No amount of effort put into quality initiatives can guarantee improvement in process performance. Very often, improvements do not stick because there is no set of systems and procedures to support and embed them into the operation's day-to-day routines. 'Quality systems' are needed. A quality system is the organisational structure, responsibilities, procedures, processes and resources for implementing quality management.[8] It should cover all facets of a business's operations and processes, and define the responsibilities, procedures and processes that ensure the implementation of quality improvement. The best-known quality system is ISO 9000.

The ISO 9000 approach

The ISO 9000 series is a set of worldwide standards that establishes requirements for companies' quality management systems. It provides a framework for quality assurance. Originally its purpose was to provide an assurance to the purchasers of products or services by defining the procedures, standards and characteristics of the control system that governed the process that produced them. It takes a 'process' approach that requires operations to define and record core processes and sub-processes. ISO 9000 also stresses four other principles:

1. Quality management should be customer focused and improvement against customer standards should be documented.
2. Quality performance should be measured and analysed, and relate to products and services themselves, the processes that created them and customer satisfaction.
3. Quality management should be improvement driven.
4. Top management must demonstrate their commitment to maintaining and continually improving management systems.

The Deming Prize

The Deming Prize was instituted by the Union of Japanese Scientists and Engineers and is awarded to those companies that have successfully applied 'company-wide quality control' based upon statistical quality control. Applicants are required to submit a detailed description of quality practices. This is a significant activity in itself and some companies claim a great deal of benefit from having done so.

Case example

Quality systems only work if you stick to them[9]

When passengers on the Hakata-to-Tokyo express, one of Japan's famous bullet trains, noticed a burning smell and an unusual sound, they were ordered off the train. The cause turned out to be cracks in the chassis and it marked the latest episode in a long line of quality scandals that had rocked the country, and caused serious reputational risk to the world's image of 'Japanese quality'. The previous months had seen public admissions by some of Japan's most prestigious names, including Kobe Steel, Mitsubishi Materials, Nissan Motor and Subaru, that their quality tests had been falsified or the results had been fabricated – all to sell products of a lower quality than officially stated. Quality systems had been in place, but often ignored. Quality records had been doctored on materials that had been shipped to make a wide range of products, including the Boeing 787 Dreamliner, nuclear plants and space rockets. It was the shock announcement from Kobe Steel that focused the world's attention on the problem. It confessed that 'improper conduct' had led to the falsification of data relating to 19,300 tonnes of aluminium sheets and poles, 19,400 aluminium components, 2,200 tonnes of copper products and an unspecified amount of iron powder that was supplied to over 200 customers. All these items had been falsely certified as conforming to specifications concerning properties such as tensile strength. Kobe admitted that for up to 10 years its employees had falsified quality checks on tens of thousands of tonnes of metal products, including the aluminium used by Boeing to make the parts that held the 787 together. However, Boeing made clear that it had been conducting comprehensive inspections and analysis of affected shipments since it was told about Kobe Steel's data falsification. *'Nothing in our review to date leads us to conclude that this issue presents a safety concern, and we will continue to work diligently with our suppliers to complete our investigation,'* it said. And, notwithstanding the falsification, no deaths or accidents seem to have resulted from the under-specification products. Also, officials at Japan's Ministry of Economy, Trade and Industry said Kobe Steel's products did not fall below industry minimum standards. It was, it said, an issue between Kobe Steel and its customers because it had not met specifications that they had demanded.

Nevertheless, Kobe Steel demoted three executives from the aluminium and copper divisions, who had been aware of the data tampering. Two executives had apparently known of the falsification problems for eight years. The company said that they were relieved from their duties and reassigned to lower-ranking roles. Also, the government-backed Japan Industrial Certification was revoked at one of its factories, owing to 'improper quality management'. But it was the nationwide inquest as to why so many problems had occurred in so many Japanese companies that was followed closely by quality professionals. Some commentators claimed that it was because of increased pressure to produce profits. When Toray Industries, a chemical company, disclosed data falsification on tyre material, Akihiro Nikkaku, its President, Chief Executive and Chief Operating Officer, blamed the 'pressure to meet productivity targets'. Other observers pointed to Japanese corporate culture and the reluctance of middle managers to communicate quality mistakes to their superiors' attention. Yet others say that the reason that so many scandals have emerged relatively recently is that, among younger employees, revealing bad practice has become more acceptable. Moreover, social media has provided a forum and an environment for whistle-blowing and airing such grievances, which previously did not exist.

RichLegg/E+/Getty Images

The EFQM Excellence Model

Over 20 years ago, Western European companies formed the European Foundation for Quality Management (EFQM). Since then, the importance of quality excellence has become far more widely accepted. While there are many frameworks and tools that are used to establish quality in organisations, the EFQM Excellence Model provides an holistic view of an organisation and it can be used to determine how these different methods fit together and complement each other. The model is an overarching framework for developing sustainable quality excellence. Excellent organisations achieve and sustain outstanding levels of performance that meet or exceed the expectations of all their stakeholders.

The model is based on the idea that it is important to understand the cause-and-effect relationships between what an organisation does (what it terms 'the enablers') and the results it achieves. The EFQM Excellence Model is shown in Figure 13.9. There are five enablers:

1. ***Leadership*** – that looks to the future, acts as a role model for values and ethics, inspires trust, is flexible, enables anticipation and so can react in a timely manner.

2. ***Strategy*** – that implements the organisation's mission and vision by developing and deploying a stakeholder-focused strategy.

3. ***People*** – organisations should value their people, creating a culture that allows mutually beneficial achievement of both organisational and personal goals, develops the capabilities of people, promotes fairness and equality, cares for, communicates, rewards and recognises people in a way that motivates and builds commitment.

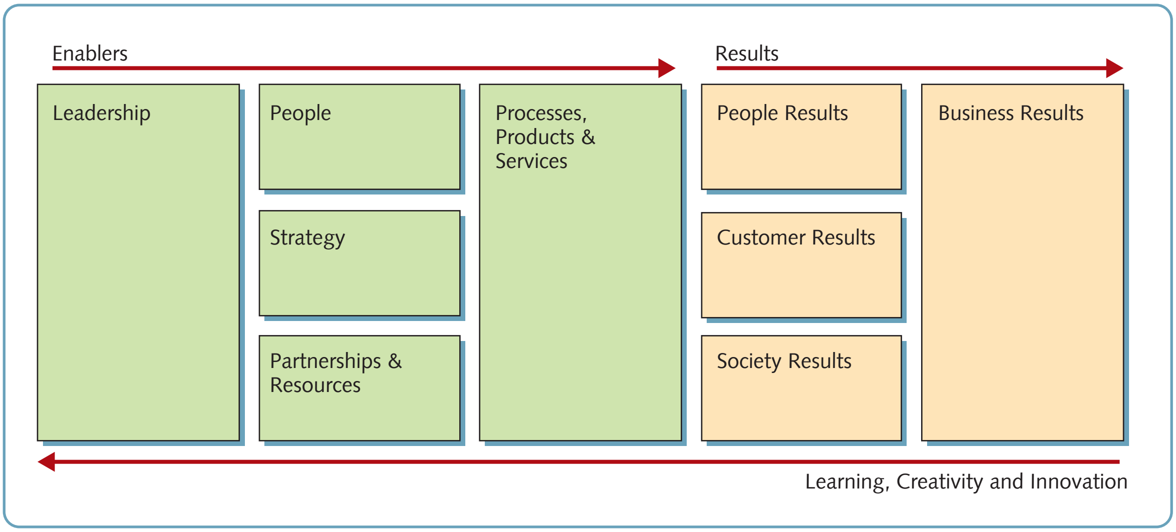


Figure 13.9 **The EFQM Excellence Model**
Source: reproduced with the permission of the EFQM

4. ***Partnerships and resources*** – organisations should plan and manage external partnerships, suppliers and internal resources in order to support strategy and policies, and the effective operation of processes.
5. ***Processes, products and services*** – organisations should design, manage and improve processes to ensure value for customers and other stakeholders.

Results are assessed using four criteria. They are:

1. ***Customer results*** – meeting or exceeding the needs and expectations of customers.
2. ***People results*** – meeting or exceeding the needs and expectations of employees.
3. ***Society results*** – achieving and sustaining results that meet or exceed the needs and expectations of the relevant stakeholders within society.
4. ***Business results*** – achieving and sustaining results that meet or exceed the needs and expectations of business stakeholders.

Critical commentary

- Quality management has been one of the hottest topics in operations management and one of the most controversial. Much of the debate has centred on the people-focus of quality management, especially the rhetoric of employee empowerment central to several modern approaches to quality. In many cases, it can be little more than an increase in employee discretion over minor details of their working practice. Some industrial relations academics argue that TQM rarely affects the fundamental imbalance between managerial control and employees' influence over organisational direction. They believe that there is little evidence that employee influence over corporate decisions that affect them has been, or can ever be, enhanced through contemporary configuration of involvement. While involvement might increase individual task discretion, or open up channels for communication, it is not designed to offer opportunities for employees to gain or consolidate control over the broader environment in which their work is located.[10]

- Other criticisms concern the appropriateness of some mechanisms such as service-level agreements (SLAs). Some see the strength of SLAs as the degree of formality they bring to customer–supplier relationships, but there are also drawbacks. The first is that the 'pseudo-contractual' nature of the formal relationship can work against building partnerships. This is especially true if the SLA includes penalties for deviation from service standards. The effect can sometimes be to inhibit rather than encourage joint improvement. The second is that SLAs tend to emphasise the 'hard' and measurable aspects of performance, rather than the 'softer' but often more important aspects. So a telephone may be answered within four rings, but how the caller is treated in terms of 'friendliness' may be far more important.
- Similarly, and notwithstanding its widespread adoption (and its revision to take into account some of its perceived failings), ISO 9000 is not seen as beneficial by all authorities. Criticisms include that it is expensive and time-consuming, that the time and cost of achieving and maintaining ISO 9000 registration are excessive, and that it is too formulaic, encouraging 'management by manual'.

SUMMARY CHECKLIST

- ☐ Does everyone in the business really believe in the importance of quality, or is it just one of those things that people say without really believing it?
- ☐ Is there an accepted definition of quality used within the business?
- ☐ Do people understand that there are many different definitions and approaches to quality, and do they understand why the business has chosen its own particular approach?
- ☐ Do all parts of the organisation understand their contribution to maintaining and improving quality?
- ☐ Are service-level agreements used to establish concepts of internal customer service?
- ☐ Is some form of gap model used to diagnose quality problems?
- ☐ Is quality defined in terms of a series of quality characteristics?
- ☐ Is quality measured using all relevant quality characteristics?
- ☐ Is the cost of quality measured?
- ☐ Are quality costs categorised as prevention, appraisal, internal failure and external failure costs?
- ☐ Is quality adequately controlled?
- ☐ Has the idea of statistical process control (SPC) been explored as a mechanism for controlling quality?
- ☐ Do individual processes have any idea of their own variability of quality performance?
- ☐ Have quality systems been explored, such as ISO 9000 and the EFQM Excellence Model?

Case study

Rapposcience Labs

'*There is no doubt that it was a disaster for the laboratory. It was the first time that a client had withdrawn from a contract so soon, and it was our fault entirely. It was also a disaster for Vincent. I feel sorry for him. I had known him for years. He was a good guy with seemingly unlimited energy and a host of good ideas. But in the end, he had to go.*' (Petra Reemer, Chief Scientist, Rapposcience Labs)

Petra Reemer was talking about her predecessor, Vincent De Smet, who was in charge of the laboratories (simply known internally as 'the lab') when one of their larger clients, MGQ Services, an extraction services firm, had exercised its right to withdraw from a commercial contract with Rapposcience for '*persistent and significant failure to comply with testing and analytical performance*'. This came as a shock to the lab because, although they were aware that their performance had not been entirely satisfactory, MGQ had not formally complained about the lab's performance. MGQ's withdrawal not only a created a hole in the lab's revenue projections, it also attracted enough negative publicity in the industry for the lab's private equity owners, Brighthorpe Holdings, to replace Vincent De Smet with Petra Reemer. With a background in analytical and industrial forensic testing, Petra started the job of rescuing the lab's reputation.

Rapposcience labs

Rapposcience Labs was located at Beveren, near Antwerp in Belgium. In the past, it had been one of the most reputable labs for analysing mineral deposits, soil and mixed inert and biological samples for a number of clients, mainly from extraction (mining), oil and gas, and public environmental agencies. It employed 47 staff, almost all with a science or technical background, the majority in testing and analysis roles together with some in administrative and sales roles. Up until the MGQ 'disaster', Brighthorpe had adopted a 'hands off' policy towards how the lab was run. That changed after De Smet's replacement, and Petra Reemer had been given the clear message that she must turn Rapposcience around, or its future would be bleak. '*We lost the MGQ contract in February. Ironically, the previous 12 months had brought in record levels of business for the lab. Yet it was business won by undercutting rivals on price. In fact, with hindsight, it is obvious that we had been running at a marginal loss all that year. I arrived in March, and I have spent the last month doing my best to reassure our remaining clients that they can still trust us to deliver a timely and trustworthy service. Unfortunately, a couple of contracts were up for renewal at that time and, regretfully, we lost them. We are now running at what looks like a sustained loss for the first time in our history*.' (Petra Reemer)

The Rapposcience laboratory process

The laboratory divided its activities into four phases of what it called its 'testing cycle'. These were pre-contract, field operations, analytics and post-analytics. Table 13.6 summarises these phases

Pre-contract occurred at the start of the contract, and involved agreeing with the client the exact specification of the service to be provided. This usually included the range of sample specifications, how they would be delivered to the lab, the nature of the report that would be prepared, and the contracted performance in terms of analytical accuracy (which indicates the veracity of the analysis), precision (which indicates the reproducibility of the analysis) and the timeliness of the report. Laboratory errors had a reported frequency of between 0.012 per cent and 0.6 per cent. Although not large in itself, errors can have a huge impact on clients' decision-making, as 60–70 per cent of their operational and investment decisions were made on the basis of laboratory tests.

Field operations was the responsibility of the client, but the lab often supplied the containers used for the samples, and instructions for taking and packaging the samples. Some clients also insisted on more detailed sampling protocols for their field technicians, including training packs.

The analytics phase included all the testing within the lab itself. This would vary depending on the nature of the tests and the procedures specified in the contract. Generally though, all testing followed three stages; sample preparation, pre-analysis treatment and analysis (see Figure 13.10).

Table 13.6 **The testing cycle**

Phase of testing cycle			
Pre-contract	***Field operations***	***Analytics***	***Post-analytics***
• Sample specification • Delivery to lab agreed • Report outline	• Sampling protocols (including training pack) • Containing and recording • Couriering	• Sample preparation • Pre-analysis treatment • Analysis	• Report generated • Data recording

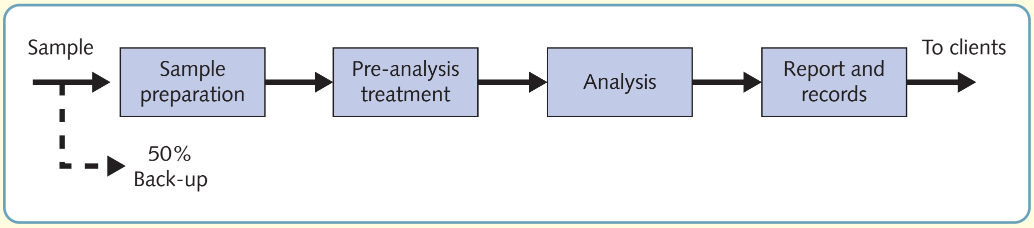


Figure 13.10 The testing phases of Rapposcience Labs' laboratory process

One of the first modifications to the process came when Vincent had decided to split the sample into two parts before it was tested. Almost always there was sufficient material to be able to do this, and the advantage was that, if the testing proved inconclusive, or some performance indicators were outside the permitted range, the tests could be repeated. Performance indicators demonstrated whether the analytical process was behaving as planned, if it had revealed a statistical anomaly that required investigation, or when a test had failed. Most contracts specified a particular confidence level for the results (usually 99.5 per cent), but any small error or contamination in the testing procedure could reduce the confidence level. If this happened, the 'back-up' sample could be tested. However, this almost certainly meant that the lab would not be able to meet its promised report delivery time.

The post-analytics phase consisted of preparing the results of the analysis for the client. This was usually a simple report describing the composition of the sample, but some clients also required a more detailed comparative report, where sample data were compared with previous sample readings. Even if such comparative reporting was not required, the lab recorded all sample data.

Initiatives during the De Smet period

Petra Reemer was not unsympathetic to what Vincent De Smet had been trying to do at Rapposcience. Not only had Vincent tried to introduce some worthwhile reforms to the lab's operating procedures, he was labouring under pressure to increase the profitability of the operation. '*I think that Vincent had been trying to increase the volume of business while keeping staffing levels the same. Presumably he figured that increased revenue with costs held down would equal healthy profitability. He also complicated things by introducing a number of initiatives, all at more or less the same time.*'

One of Vincent's initiatives had been his decision to split the sample into two parts before it was tested. He had done this as a 'failsafe' in case there were problems during the analysis phase and the tests had to be repeated. The response of the lab's technicians to this move had been mixed. Some felt that it was a sensible move that reduced the chances of recording a 'failed through insufficient material' result. Although this did not happen often, it was at best embarrassing to the lab and at worst extremely irritating for the client. Others felt that, because there was the possibility of retesting a sample, there was a tendency to take less care and 'adopt testing shortcuts' because the consequences of testing errors were less serious.

Another of Vincent's innovations had been the introduction of limited statistical process control (SPC). Although the lab had always recorded measures of its analytical performance, it had not formally examined its analytics process performance in any systematic manner. It was the MGQ contract that Vincent won (and lost) that prompted the lab to take the potential of SPC seriously. During the pre-contract phase, MGQ had insisted on its use during all testing on its samples, together with periodic SPC summaries being submitted. Vincent had invested in a 'smart laboratory' IT system that was advertised as being able to automate the data management and statistical processes in the lab. However, almost a year after its partial introduction, the consensus in the lab was that it had not been a success. '*It was just too sophisticated for us*', said Petra Reemer, '*we were trying to run before we could walk*'.

The final initiative instituted during Vincent's time as Chief Scientist was an enhanced set of reporting protocols. '*It wasn't a bad idea actually*', admitted Petra Reemer, '*we already prepared more extensive reports for some clients, so we had the expertise to interpret their test results and advise them on their sampling processes and how they might interpret results. In other words, we have expertise that can add real value for our clients, so why not use it to enhance our quality of service? The problem when Vincent introduced the idea was that he tried to push it as a sales promotion tool. Clients were inclined to dismiss the potential of enhanced reporting because they thought that we were simply trying to get more money out of them.*'

Getting back to basics

Petra had taken over from Vincent in March. After three or four weeks talking with all the staff in the lab, she felt she was ready to shape her plans for the lab's future. She was convinced that the lab had to understand what really mattered to clients and then do everything to improve their performance in a way that would have an impact on the quality of service they were providing. Unfortunately, she was also facing pressure from Brighthorpe, the lab's owners, to cut costs. '*I persuaded them to give me time to restore our reputation. We would find it difficult to do that if we were shedding staff at the same time. Not only would it send the wrong message to the market, it would make it difficult to improve the way we do things. Having said that, we decided not to replace any staff who left* *the lab of their own volition. We also delayed any non-essential expenditure. The main objective was to survive long enough to get back to the basics of how we could serve clients better*.'

Her first action was to look at how SPC had been used in the lab since it had been introduced. She talked with the chief field engineer at MGQ, who had approved the initial contract that the lab had lost, and who had also insisted on them using SPC. What he said gave Petra much to think about. '*I kind of knew that, when we insisted on Rapposcience using SPC that they really didn't understand what it was all about. They were simply doing it because it was what the client wanted. Their culture said, "If the samples are returned as the specification, then it's OK; if not, then as long as it doesn't happen too often, well that's OK also". They just didn't get that by seeing their process charts, it enabled us to see more or less exactly what was happening right inside their processes. I take some of the blame myself. I should have made sure that they fully understood why we were so keen for them to use SPC. It was for them to help themselves by improving their process performance. It wasn't just a whim on our part*.' (Chief Field Engineer, MGQ)

The first thing Petra did was to hold a series of meetings, first with the supervisors in each department, then with everyone in each department. She was mainly listening to their experiences of using the SPC system that Vincent had imposed, but her secondary motive was to try to judge how much they understood about the fundamentals of SPC. The answer seemed to be 'not a lot'. They were all used to using quite sophisticated statistics within their testing procedures, but not for controlling the performance of the processes themselves. Petra reflected on this: '*I guess it's because the statistics that our technicians use every day are essentially static. They deal with the probability of certain elements or contaminants being present in a single sample. SPC deals with dynamic probabilities – time series in effect – that show whether process behaviour is changing. However, the positive outcome from these meetings was that staff had little problem understanding the basic concepts of SPC, when they were explained. They were not frightened by the maths*.'

Petra realised that, in fact, the biggest problem was attitudinal. '*We had been working for a year with the attitude that testing productivity was paramount. Don't waste time. Get as many tests done as possible every day. It took time to move to an attitude that stressed error-free testing. What was the point of carrying on with testing when the processes themselves were "out of control". They would only have to be repeated, wasting everyone's time. It may be counter-intuitive, but being slow but methodical, and checking the process regularly, can actually increase effective productivity*.' With the agreement of her staff, Petra devised a set of 'check rules'. These were reference values for all the major procedures in the sample preparation, pre-analysis and analysis stages, which indicated that test results at any stage, although within the limits that indicated a reliable result, were close to those limits. If results violated these 'check rules', the test would be suspended and the sample investigated before it was allowed to progress. Petra had three reasons for instituting the 'check rules'. First, it prevented effort being wasted on samples that could be compromised. Second, it stressed the importance of trying to investigate the root causes of any problems with the process. Third, it emphasised the importance of the lab's processes in determining their quality of service to customers, and therefore to the lab's profitability and survival.

The 'root cause' programme

By September the lab's process performance had improved to the point where the number of samples that failed the reliability test had almost halved, and the number of late reports had fallen by over a third. But Petra believed that further improvements were possible. '*The most significant change is in the lab's culture. Before, staff were simply going through the motions. They were not deliberately being careless, but they were not really digging beneath what they were doing, they were not building their process knowledge. If asked, they would tell you what they were doing rather than why they were doing it. Now there is genuine curiosity about how testing procedures could be made better*.'

Petra wanted to use the staff's new-found interest in the process to make further improvements through what she called the 'root cause' initiative. As the name implies, this was a push to discover what was causing problems in testing. The data collected from those occasions where the

Table 13.7 Root cause by phase of the testing process

	Phase of testing			
	Sample preparation (62% of total errors)	*Pre-analysis (19% of total errors)*	*Analysis (15% of total errors)*	*Report and record (4% of total errors)*
Types of root cause	Mislabelled sample (F) (24%) Badly contained (F) (18%) Preparation error (8%) Request error (F) (6%) Insufficient material (F) (3%) Damaged sample (F) (3%)	Reagent error (6%) Contamination (5%) Spillage (5%) Process violation (3%)	Calibration error (6%) Process violation (4%) In-test calculation (3%) Contamination (2%)	Reading error (2%) Interpretation error (1%) Missing data (1%)

(F) = root cause in the field (client's responsibility) check rules had been invoked provided valuable information, which was further supplemented by individual investigations by 'root cause teams' in each department. Petra, with the support of supervisors in each department, had encouraged the formation of these teams, but not made them compulsory. However, most staff elected to become 'root cause team' members.

By the end of October, Petra was in a position to consolidate all the data on the root causes of all the occasions when an error of some sort had occurred in the lab's processes. This included any defect, from ordering tests to reporting and interpretation of the results. Table 13.7 shows the root causes.

What was interesting to Petra was the dominance of errors with a root cause outside the lab. The data indicated that more than half of all errors were outside the scope of the lab's responsibility. '*This shouldn't lead us into any form of complacency. We can still do a lot to tackle the errors in the phases of the process for which we are clearly responsible. Basic laboratory procedure, such as choosing the incorrect reagent, violating process rule, or allowing contamination, should not be happening. Also, I suspect that we are actually committing more "errors" in the "report and record" phase than it seems. Errors in testing are more obvious, but reporting is not always right or wrong. There are probably opportunities to enhance our service to clients that we are missing. You could class them as just as much of an "error" as a contaminated sample*.'

Questions

1. In hindsight, what were Vincent's mistakes in running the lab?
2. How did Petra's approach differ, and why was it more successful?
3. Is a 'missed opportunity' in the report and record stage as much of an error as a contaminated sample, as Petra suggests?
4. What do you suggest that Petra does next to improve process quality further?

Applying the principles

Some of these exercises can be answered by reading the chapter. Others will require some general knowledge of business activity and some might require an element of investigation. **Model answers for the first two questions are available on the student companion website**. *For model answers to the other questions in this section, please ask your tutor/lecturer.*

1. Human error is a significant source of quality problems. Think through the times that you have (with hindsight) made an error and answer the following questions:

a) How do you think that human error causes quality problems?

b) What could one do to minimise human error?

2. The owner of a small wedding photography business realises that the market is changing. '*I used to take a few photos during the wedding ceremony and then formal group shots outside. It rarely took more than two hours. Around 30 photos would go in a standard wedding album. You had to get the photos right, that was really the only thing I was judged on. Now it's different. I spend all day at a wedding, and sometimes late into the evening. You're almost like another guest. All the guests at the wedding now are important. You have to get the best photos while being as discreet as possible. Clients judge you on both the pictures and the way you interact with everyone on the day. The product has changed too. Clients receive a memory stick with around 500 photos and a choice of 10 albums. I also offer photo books with greater customisation. I can offer albums with items such as invitations, confetti and menus; and individual paintings created from photographs. Obviously, I would have to outsource the paintings. Wedding guests can order photos and related products online. My anxiety is that advertising this service at the wedding will be seen as being too commercial. We have a high level of demand in summer, with weekends booked up two years in advance. I may take on additional photographers during busy periods, but the best ones are busy themselves. Also, the business is about client relations and that's hard to replicate. I often offer clients advice on such things as locations, bands, caterers and florists. Wedding planning is clearly an area that could be profitable to the business. Another option is to move beyond weddings into other areas, such as school photos, birthdays, celebrations or studio work*.'

a) How has the business changed over time?

b) What do you think are the key quality challenges facing the business?

c) What do you think should be done to ensure the business maintains high quality levels in the future?

3. Ryanair is the best-known budget airline in Europe, focusing on popular routes and very low operating costs. For years, the boss of the airline Michael O'Leary's policy on customer service was clear. '*Our customer service*', he said, '*is about the most well defined in the world. We guarantee to give you the lowest air fare. You get a safe flight. You get a normally on-time flight. That's the package. We don't, and won't, give you anything more. Are we going to say sorry for our lack of customer service? Absolutely not. If a plane is cancelled, will we put you up in a hotel overnight? Absolutely not. If a plane is delayed, will we give you a voucher for a restaurant? Absolutely not*.' However, the bad publicity eventually prompted a limited rethink by the company. After a drop in their hitherto rapid profit growth, some shareholder concern and Ryanair being voted the worst of Britain's 100 biggest brands by readers of the consumer magazine *Which?*, the company announced that it was to reform its '*abrupt culture, and try to eliminate things that unnecessarily annoy customers*'. Included in these annoying practices were fines for small luggage-size transgressions and an unpopular €70 fee for issuing boarding passes at the airport rather than printing them out at home (it was lowered to €10). Yet Ryanair insisted that such charges were not money-spinning schemes, but were designed to encourage operational efficiency that kept fares low; in fact fewer than 10 passengers a day had to pay for forgotten boarding passes. What does this example tell us about the trade-off between service quality and cost?

4. Understanding type I and type II errors is essential for surgeons' quality planning. In appendectomy operations, for example, removal of the appendix is necessary because of the risk of it bursting and causing potentially fatal poisoning of the blood. The surgical procedure is a relatively simple operation but there is always a small risk with any invasive surgery. It is also expensive; in the US around $4,500 per operation. Unfortunately, appendicitis is difficult to diagnose; diagnosis is only 10 per cent accurate. However, a new technique claims to be able to identify 100 per cent of true appendicitis cases prior to surgery. The new technique costs less than $250, which means that one single avoided surgery pays for around 20 tests.

a) How does this new test change the likelihood of type I and type II errors?

b) Why is this important?

5. Tea and Sympathy is a British restaurant and café in the heart of New York's West Village. It is tiny, with around a dozen tables packed into an area little bigger than the average British sitting room. Expatriate Brits, native New Yorkers and celebrities queue to get in. It has become famous for the unusual nature of its service. '*Everyone is treated in the same way*', says Nicky Perry who runs it. '*We have a firm policy that we don't take any shit.*' This robust attitude to the treatment of customers is reinforced by 'Nicky's Rules', which are printed on the menu:

- Be pleasant to the waitresses – remember Tea and Sympathy girls are always right.
- You will have to wait outside the restaurant until your entire party is present: no exceptions.
- Occasionally, you may be asked to change tables so that we can accommodate all of you.
- If we don't need the table you may stay all day, but if people are waiting it's time to naff off.

If customers object they are thrown out. Nicky says that she has had to train 'her girls' to toughen up. '*I've taught them that when people cross the line they can tear their throats out as far as I'm concerned. What we've discovered over the years is that if you are really sweet, people see it as a weakness*.' People get thrown out of the restaurant about twice a week.

a) Why do you think 'Nicky's Rules' help to make the Tea and Sympathy operation more efficient?

b) The restaurant's approach to quality of service seems very different to most restaurants. Why do you think it seems to work here?

6. Four Seasons Hotels is a chain of very 'upmarket' hotels famed for their quality of service. From its inception the group has had the same guiding principle, '*to make the quality of our service our competitive advantage*'. The company has what it calls its Golden Rule: 'Do to others (guests and staff) as you would wish others to do to you'. It is a simple rule, but it guides the whole organisation's approach to quality. All employees are empowered to use their creativity and judgement in delivering exceptional service and making their own decisions to enhance their guests' stay. For example, one morning an employee noticed that a guest had a flat tyre on their car and decided of his own accord to change it for them, which was very much appreciated by the guest. Managers have a focus on recognising these successes and publicly praise and celebrate all individuals who deliver these thoughtful touches. Their objective is to exceed guest expectations. They have created an in-house database that is used to record all guest feedback (whether positive or negative).

a) The company has what it calls its Golden Rule: 'Do to others (guests and staff) as you would wish others to do to you'. Why is this important in ensuring high-quality service?

b) What do you think the hotels' guests expect from their stay?

c) How does staff using their own initiative contribute to quality of service?

Notes on chapter

1 Qualinet (2012) 'White Paper on Definitions of Quality of Experience', Qualinet.

2 The information on which this example is based is taken from: Giugliano, F. (2015) 'Bank of England moves to stamp out fat finger errors', *Financial Times*, 14 June; Economist (2013) 'Overtired, and overdrawn', *Economist* print edition, 15 June; Wilson, H. (2014) 'Fat-fingered trader sets Tokyo alarms ringing', *The Times*, 2 October.

3 Zeithaml, V.A., Parasuraman, A. and Berry, L.L. (1990) *Delivering Quality Service: Balancing customer perceptions and expectations*, Free Press.

4 The information on which this example is based is taken from: Markillie, P. (2011) 'They trash cars, don't they?', *Intelligent Life Magazine*, Summer.

5 Plunkett, J.J. and Dale, B.S. (1987) 'A review of the literature in quality-related costs', *The International Journal of Quality & Reliability Management*, 4 (1), pp. 40–52.

6 The information on which this example is based is taken from: Walne, T. (2019) 'Want to exchange a jar of coins for notes?', This is Money website, 24 August, https://www.thisismoney.co.uk/money/betterbanking/article-7390729/Want-notes-not-coins-Banks-dont-care-cash-investigation-finds.html [accessed 1 October 2020]; Schubber, K. (2016) 'The Metro Bank coin caper', *Financial Times*, 2 June; Morgan, R. (2016) 'TD Bank dumps its faulty coin-counting machines', *New York Post*, 19 May.

7 The information on which this example is based is taken from: Wheatley, M. (2010) 'Filling time on the production line', *Engineering and Technology*, 8 November.

8 Dale, B.G., Bamford, D. and van der Wiele, T. (eds) (2016) *Managing Quality: An essential guide and resource gateway*, 6th edition, John Wiley & Sons.

9 The information on which this example is based is taken from: Wells, P. and Lewis, L. (2018) 'Japan Inc: A corporate culture on trial after scandals', *Financial Times*, 3 January; Parry, R.L. (2017) 'Japan's failed corporate culture at root of Kobe Steel scandal', *The Times*, 16 October; Economist (2017) 'Kobe Steel admits falsifying data on 20,000 tonnes of metal', *Economist* print edition, 12 October; Wells, P. (2017) 'Kobe Steel demotes three executives over cheating scandal', *Financial Times*, 21 December; Wells, P. and Terazono, E. (2017) 'Five questions on Kobe Steel and quality controls', *Financial Times*, 11 October.

10 For an example of this type of critique, see Hyman, J. and Mason, B. (1995) *Managing Employee Involvement and Participation*, Sage.

Taking it further

Crosby, P.B. ***(1979) Quality is Free: The art of making quality certain, McGraw-Hill.*** *One of the gurus. It had a huge impact in its day. Read it if you want to know what all the fuss was about.*

Dale, B.G., Bamford, D. and van der Wiele, T. ***(2016) Managing Quality, 6th edition, Wiley.*** *This is the fifth edition of a book that has long been one of the best-respected texts in the area. A comprehensive and balanced guide.*

Feigenbaum, A.V. ***(1986) Total Quality Control, McGraw-Hill.*** *A more comprehensive book than those by some of the other quality gurus.*

Goetsch, D.L. and Davis, S. ***(2013) Quality Management for Organizational Excellence: Introduction to total quality, 7th edition, Pearson.*** *Up-to-date account of the topic*.

Hoyle, D. ***(2006) Quality Management Essentials, Butterworth-Heinemann.*** *A practical guide*.

Oakland, J.S. ***(2014) Total Quality Management and Operational Excellence: Text with cases, Pearson.*** *The latest text from one of the gurus of quality*.

Pande, P.S., Neuman, R.P. and Kavanagh, R.R. ***(2000) The Six Sigma Way, McGraw-Hill Education.*** *There are many books written by consultants for practising managers on the now-fashionable Six Sigma approach. This one is readable and informative*.

Supplement to Chapter 13
Statistical process control (SPC)

The purpose of statistical process control (SPC) is both to control process performance, keeping it within acceptable limits, and to improve process performance by reducing the variation in performance from its target level. It does this by applying statistical techniques to understand the nature of performance variation over time. For those who are anxious about the 'statistical' part of SPC, don't be. Essentially SPC is based on principles that are both practical and intuitive. The statistical element is there to help rather than complicate quality decisions.

Variation in process performance

The core instrument in SPC is the control chart. (These were explained earlier in the chapter.) They are an illustration of the dynamic performance of a process, measuring how some aspect of process performance varies over time. All processes vary to some extent. No machine will give precisely the same result each time it is used. All materials vary a little. People in the process differ marginally in how they do things each time they perform a task. Given this, it is not surprising that any measure of performance quality (whether attribute or variable) will also vary. Variations that derive from these *normal* or *common* causes of variation can never be entirely eliminated (although they can be reduced).

For example, at a call centre for a utility company, customer care operatives answer queries over accounts, service visits, and so on. The length of time of each call will vary depending on the nature of the enquiry and the customer's needs. There will be some variation around an average call time. When the process of answering and responding to customer enquiries is stable, the computer system that intercepts and allocates calls to customer care operatives could be asked to randomly sample the length of each call. As this data is built up, the histogram showing call times could develop as is shown in Figure 13.11. The first calls could lie anywhere within the natural variation of the process, but are more likely to be close to the average call length (Figure 13.11(a)). As more calls are measured, they would clearly show a tendency to be close to the process average (see Figure 13.11(b) and (c)). Eventually the data will show a smooth histogram that can be drawn into a smoother distribution that will indicate the underlying process variation (the distribution shown in Figure 13.11(f)).

Often this type of variation can be described by a normal distribution. (Even if this raw data does not conform to a normal distribution, it can be manipulated to approximate to one by using sampling, see later.) It is a characteristic of normal distributions that 99.7 per cent of the measures will lie within ± 3 standard deviations of the distribution (standard deviation is a measure of how widely the distribution is spread or *dispersed*).

The central limit theorem

Not all processes will vary in their performance according to a normal distribution. However, if a sample is taken from any type of distribution, the distribution of the average of the sample (sample mean) will approximate to a normal distribution. For example, there is an equal probability of any number between one and six being thrown on a six-sided, unweighted die. The distribution is rectangular with an average of 3.5, as shown in Figure 13.12(a). But if a die is thrown (say) six times repeatedly and the average of the six throws calculated, the sample average will also be 3.5, but the standard deviation of the distribution will be the standard deviation

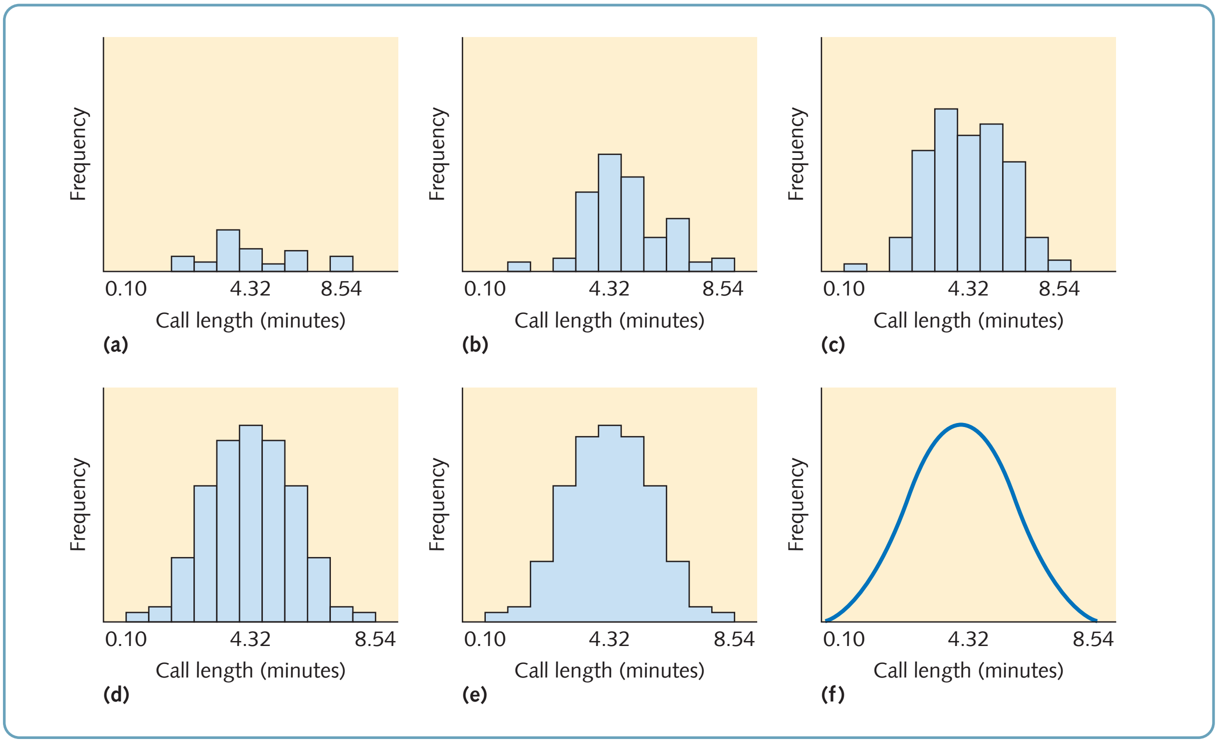


Figure 13.11 The natural variation of call times in a call centre can be described by a normal distribution

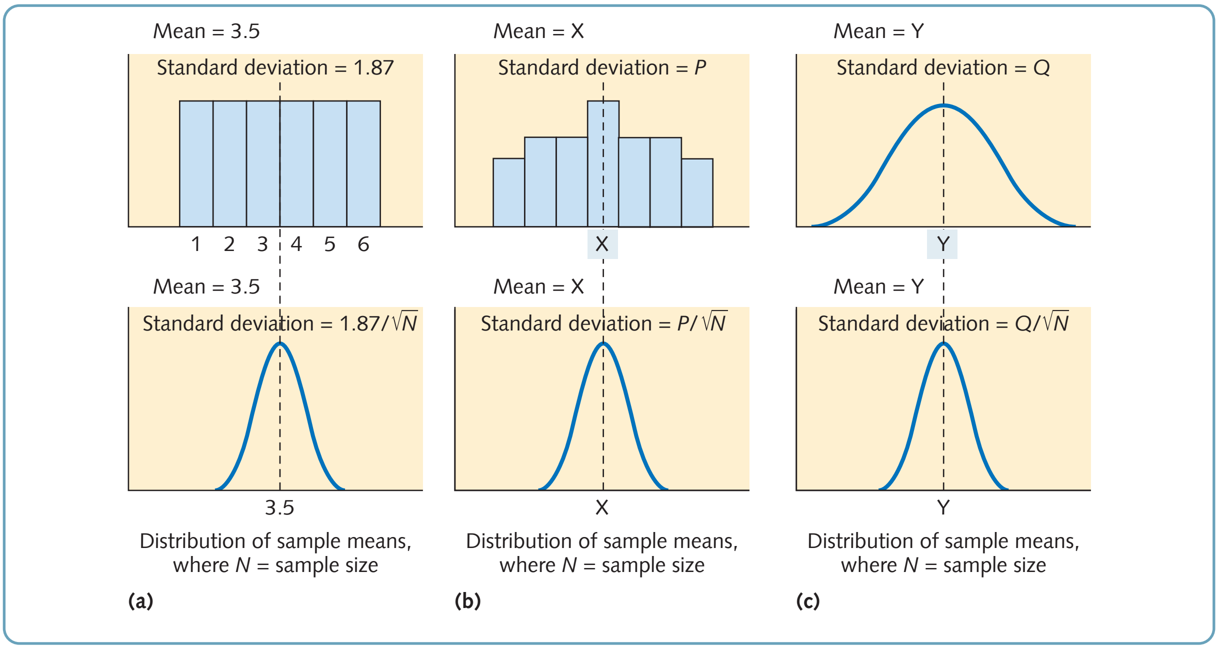


Figure 13.12 The distribution of sample means (averages) from any distribution will approximate to a normal distribution

of the original rectangular distribution divided by the square of the sample size. More significantly, the shape of the distribution will be close to normal and so can be treated the same way as a normal distribution. This becomes important when control limits are calculated (see later).

Is the process 'in control'?

Not all variation in process performance is the result of common causes. There may be something wrong with the process that is assignable to an abnormal and preventable cause. Machinery may have worn or been set up badly. An untrained person may not be following the prescribed procedure for the process. The causes of such variation are called *assignable* or *abnormal* causes. The question for operations management is whether the results from any particular sample, when plotted on the control chart, simply represent the variation due to *common* causes or due to some specific and correctable *assignable* cause. Figure 13.13(a), for example, shows the control chart for the average call length of samples of customer calls in a utility's call centre. Like any process, the results vary, but the last three points seem to be lower than usual. The question is whether this is natural variation or the symptom of some more serious cause. Is the variation the result of common causes, or does it indicate assignable causes (something abnormal) occurring in the process?

To help make this decision, control limits can be added to the control charts that indicate the expected extent of 'common-cause' variation. If any points lie outside these control limits, then the process can be deemed *out of control* in the sense that variation is likely to be due to assignable causes. These can be set in a statistically revealing manner based on the probability that the mean of a particular sample will differ by more than a set amount from the mean of the population from which it is taken. Figure 13.13(b) shows the same control chart as Figure 13.13(a) with the addition of control limits put at ±3 standard deviations (of the population of sample means) away from the mean of sample averages. It shows that the probability of the final point on the chart being influenced by an assignable cause is very high indeed. When the process is exhibiting behaviour that is outside its normal 'common-cause' range, it is said to be 'out of control'.

However, we cannot be absolutely certain that the process is out of control. There is a small but finite chance that the point is a rare but natural result at the tail of its distribution. Stopping the process under these circumstances would represent a type I error because the process is actually in control. Alternatively, ignoring a result which in reality is due to an assignable cause is a type II error (see Table 13.8). Control limits that are set at three standard deviations either side of the population mean are called the upper control limit (UCL) and lower control limit (LCL). There is only a 0.3 per cent chance of any sample mean falling outside these limits by chance causes (that is, a chance of a type I error of 0.3 per cent).

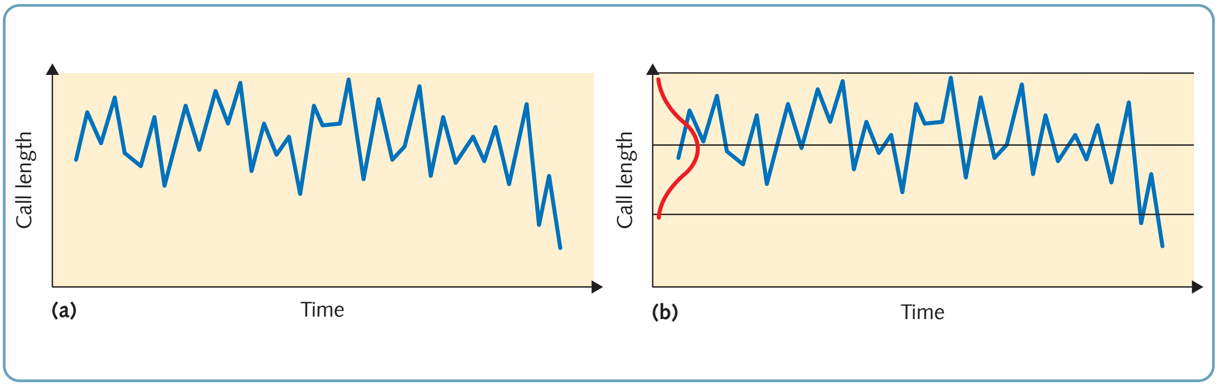


Figure 13.13 **Control chart for the average call length in a call centre**

Table 13.8 **Type I and type II errors in SPC**

		Actual process state	
		In control	*Out of control*
Decision	**Stop process**	**Type I error**	Correct decision
	Leave alone	Correct decision	**Type II error**

Process capability

Using control charts to assess whether the process is in control is an important internal benefit of SPC. An equally important question for any operations manager would be: 'Is the variation in the process performance acceptable to external customers?' The answer will depend on the acceptable range of performance that will be tolerated by the customers. This range is called the *specification range*. Returning to the 'What a giveaway' case example, if the weight of rice in the box is too small then the organisation might infringe labelling regulations; if it is too large, the organisation is 'giving away' too much of its product for free.

Process capability is a measure of the acceptability of the variation of the process. The simplest measure of capability (C_p) is given by the ratio of the specification range to the 'natural' variation of the process (i.e. ± 3 standard deviations):

$$C_p = \frac{\text{UTL} - \text{LTL}}{6s}$$

where UTL = the upper tolerance limit
LTL = the lower tolerance limit
s = the standard deviation of the process variability.

Generally, if the C_p of a process is greater than 1, it is taken to indicate that the process is 'capable', and a C_p of less than 1 indicates that the process is not 'capable', assuming that the distribution is normal (see Figure 13.14(a), (b) and (c)).

The simple C_p measure assumes that the average of the process variation is at the mid-point of the specification range. However, often the process average is offset from the specification

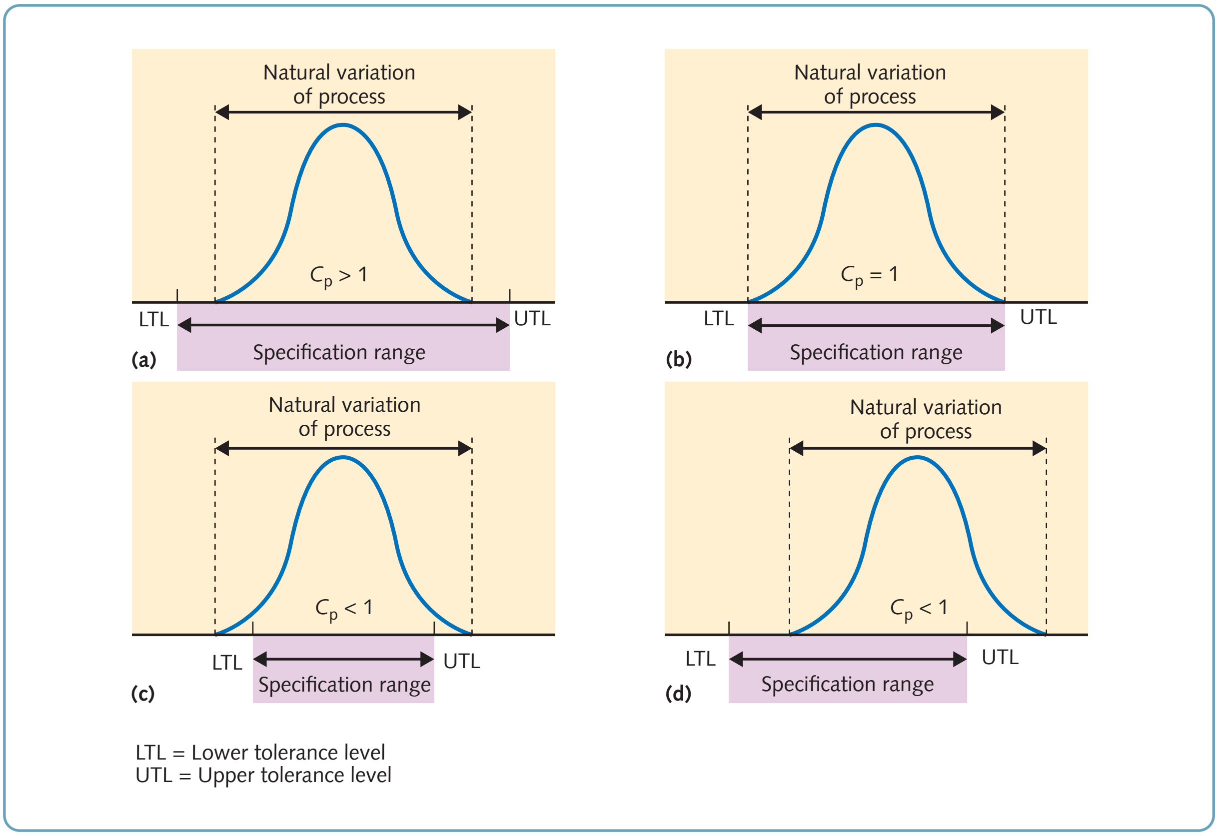


Figure 13.14 Process capability compares the natural variation of the process with the specification range that is required

range (see Figure 13.14(d)). In such cases, *one-sided* capability indices are required to understand the capability of the process:

$$\text{Upper one-sided index } C_{pu} = \frac{\text{UTL} - X}{3s}$$

$$\text{Lower one-sided index } C_{pl} = \frac{X - \text{LTL}}{3s}$$

where X = the process average.

Sometimes only the lower of the two one-sided indices for a process is used to indicate its capability (C_{pk}):

$$C_{pk} = \min(C_{pu}, C_{pl})$$

Control charts for variables

The most commonly used type of control chart employed to control variables is the $\bar{X} - R$ *chart*. In fact, this is really two charts in one. One chart is used to control the sample average or mean ($\bar{X}$). The other is used to control the variation within the sample by measuring the range (R). The range is used because it is simpler to calculate than the standard deviation of the sample.

The mean ($\bar{X}$) chart can pick up changes in the average output from the process being charted. Changes in the mean chart would suggest that the process is drifting generally away

Worked example

Box filling

In the case of the box-filling process described previously, process capability can be calculated as follows:

$$\text{Specification range} = 214 - 198 = 16 \text{ g}$$

$$\text{Natural variation of process} = 6 \times \text{standard deviation}$$

$$= 6 \times 2 = 12 \text{ g}$$

$$C_p = \text{process capability}$$

$$= \frac{\text{UTL} - \text{LTL}}{6s}$$

$$= \frac{214 - 198}{6 \times 2} = \frac{16}{12}$$

$$= 1.333$$

If the natural variation of the filling process changed to have a process average of 210 grams but the standard deviation of the process remained at 2 grams:

$$C_{pu} = \frac{214 - 210}{3 \times 2} = \frac{4}{6} = 0.666$$

$$C_{pl} = \frac{210 - 198}{3 \times 2} = \frac{12}{6} = 2.0$$

$$C_{pk} = \min(0.666, 2.0)$$

$$= 0.666$$

from its supposed process average, although the variability inherent in the process may not have changed. The range (R) chart plots the range of each sample, which is the difference between the largest and the smallest measurement in the sample. Monitoring sample range gives an indication of whether the variability of the process is changing, even when the process average remains constant.

Control limits for variables control chart

As with attributes control charts, a statistical description of how the process operates under normal conditions (when there are no assignable causes) can be used to calculate control limits. The first task in calculating the control limits is to estimate the grand average or population mean ($\bar{\bar{X}}$) and average range ($\bar{R}$) using m samples each of sample size n.

The population mean is estimated from the average of a large number (m) of sample means:

$$\bar{\bar{X}} = \frac{\bar{X}_1 + \bar{X}_2 + \ldots \bar{X}_m}{m}$$

The average range is estimated from the ranges of the large number of samples:

$$\bar{R} = \frac{R_1 + R_2 + \ldots R_m}{m}$$

The control limits for the sample means chart are:

$$\text{Upper control limit (UCL)} = \bar{\bar{X}} + A_2\bar{R}$$

$$\text{Lower control limit (LCL)} = \bar{\bar{X}} - A_2\bar{R}$$

The control limits for the range chart are:

$$\text{Upper control limit (UCL)} = D_4\bar{R}$$

$$\text{Lower control limit (LCL)} = D_3\bar{R}$$

The factors A_2, D_3 and D_4 vary with sample size and are shown in Table 13.9.

Table 13.9 Factors for the calculation of control limits

Sample size n	A_2	D_3	D_4
2	1.880	0	3.267
3	1.023	0	2.575
4	0.729	0	2.282
5	0.577	0	2.115
6	0.483	0	2.004
7	0.419	0.076	1.924
8	0.373	0.136	1.864
9	0.337	0.184	1.816
10	0.308	0.223	1.777
12	0.266	0.284	1.716
14	0.235	0.329	1.671
16	0.212	0.364	1.636
18	0.194	0.392	1.608
20	0.180	0.414	1.586
22	0.167	0.434	1.566
24	0.157	0.452	1.548

The LCL for the means chart may be negative (for example, temperature or profit may be less than zero) but it may not be negative for a range chart (or the smallest measurement in the sample would be larger than the largest). If the calculation indicates a negative LCL for a range chart then the LCL should be set to zero.

Worked example

Groupe As Maquillage

GAM (Groupe As Maquillage) is a contract cosmetics company that manufactures and packs cosmetics and perfumes for other companies. One of its plants operates a filling line that automatically fills plastic bottles with skin cream and seals the bottles with a screw-top cap. The tightness with which the screw-top cap is fixed is an important aspect of quality. If the cap is screwed on too tightly, there is a danger that it will crack; if screwed on too loosely, it might come loose. Either outcome could cause leakage. The plant had received some complaints of product leakage, possibly caused by inconsistent fixing of the screw-tops. Tightness can be measured by the amount of turning force (torque) that is required to unfasten the tops. The company decided to take samples of the bottles coming out of the filling-line process, test them for their unfastening torque and plot the results on a control chart. Several samples of four bottles were taken during a period when the process was regarded as being 'in control'.
The following data were calculated from this exercise:

$$\text{The grand average of all samples } \bar{\bar{X}} = 812 \text{ g/cm}^3$$
$$\text{The average range of the sample } \bar{R} = 6 \text{ g/cm}^3$$

Control limits for the means ($\bar{X}$) chart were calculated as follows:

$$\begin{aligned} \text{UCL} &= \bar{\bar{X}} + A_2\bar{R} \\ &= 812 + (A_2 \times 6) \end{aligned}$$

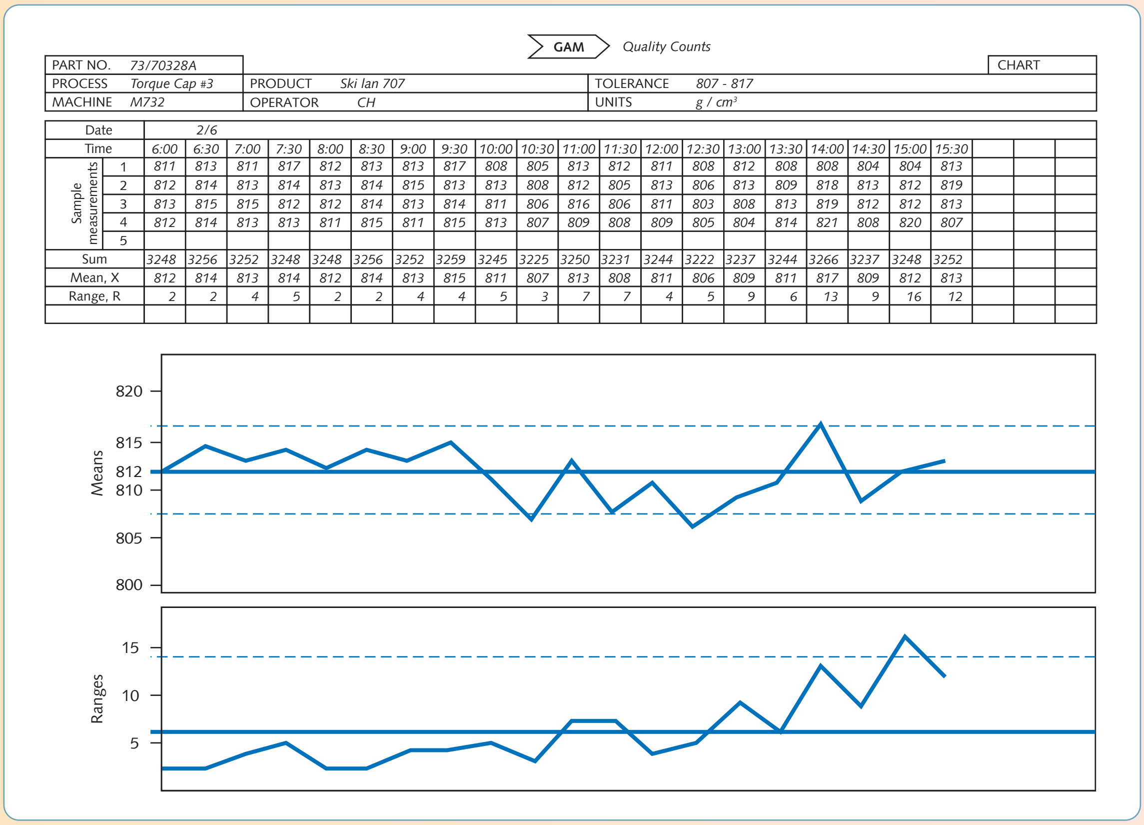
From Table 13.9, we know for a sample size of four, $A_2 = 0.729$. Thus:

$$\begin{aligned} \text{UCL} &= 812 + (0.729 \times 6) \\ &= 816.37 \\ \text{LCL} &= \bar{\bar{X}} - (A_2\bar{R}) \\ &= 812 - (0.729 \times 6) \\ &= 807.63 \end{aligned}$$

Control limits for the range chart (R) were calculated as follows:

$$\begin{aligned} \text{UCL} &= D_4 \times \bar{R} \\ &= 2.282 \times 6 \\ &= 13.69 \\ \text{LCL} &= D_3\bar{R} \\ &= 0 \times 6 \\ &= 0 \end{aligned}$$

After calculating these averages and limits for the control chart, the company regularly took samples of four bottles during production, recorded the measurements and plotted them as shown in Figure 13.15. This control chart reveals that only with difficulty could the process average be kept in control. Occasional operator interventions were required. Also, the process range was moving towards (and once exceeding) the upper control limit. The process also seemed to be becoming more variable. (After investigation it was discovered that, because of faulty maintenance of the line, skin cream was occasionally contaminating the part of the line that fitted the cap, resulting in erratic tightening of the caps.)



PART NO.	73/70328A				CHART
PROCESS	Torque Cap #3	PRODUCT	Ski lan 707	TOLERANCE	807 - 817
MACHINE	M732	OPERATOR	CH	UNITS	g / cm³

Date		2/6																			
Time		6:00	6:30	7:00	7:30	8:00	8:30	9:00	9:30	10:00	10:30	11:00	11:30	12:00	12:30	13:00	13:30	14:00	14:30	15:00	15:30
Sample measurements	1	811	813	811	817	812	813	813	817	808	805	813	812	811	808	812	808	808	804	804	813
	2	812	814	813	814	813	814	815	813	813	808	812	805	813	806	813	809	818	813	812	819
	3	813	815	815	812	812	814	813	814	811	806	816	806	811	803	808	813	819	812	812	813
	4	812	814	813	813	811	815	811	815	813	807	809	808	809	805	804	814	821	808	820	807
	5																				
Sum		3248	3256	3252	3248	3248	3256	3252	3259	3245	3225	3250	3231	3244	3222	3237	3244	3266	3237	3248	3252
Mean, X		812	814	813	814	812	814	813	815	811	807	813	808	811	806	809	811	817	809	812	813
Range, R		2	2	4	5	2	2	4	4	5	3	7	7	4	5	9	6	13	9	16	12

Figure 13.15 The completed control form for GAM's torque machine showing the mean ($\bar{X}$) and range ($\bar{R}$) charts

Control charts for attributes

Attributes have only two states – 'right' or 'wrong', for example – so the statistic calculated is the proportion of wrongs (p) in a sample. (This statistic follows a binomial distribution.) Control charts using p are called 'p-charts'. When calculating control limits, the population mean ($\bar{p}$) (the actual, normal or expected proportion of 'defectives') may not be known. Who knows, for example, the actual number of city commuters who are dissatisfied with their journey time? In such cases the population mean can be estimated from the average of the proportion of 'defectives' ($\bar{p}$), from m samples each of n items, where m should be at least 30 and n should be at least 100:

$$\bar{p} = \frac{p^1 + p^2 + p^3 \ldots p^n}{m}$$

One standard deviation can then be estimated from:

$$\sqrt{\frac{\bar{p}(1 - \bar{p})}{n}}$$

The upper and lower control limits can then be set as:

$$\text{UCL} = \bar{p} + 3 \text{ standard deviations}$$

$$\text{LCL} = \bar{p} - 3 \text{ standard deviations}$$

Of course, the LCL cannot be negative, so when it is calculated to be negative it should be rounded up to zero.

Worked example

Koops credit card company

Koops credit card company deals with many hundreds of thousands of transactions every week. One of its measures of the quality of service it gives its customers is the dependability with which it mails customers' monthly accounts. The quality standard it sets itself is that accounts should be mailed within two days of the 'nominal post date' that is specified to the customer. Every week, the company samples 1,000 customer accounts and records the percentage that was not mailed within the standard time. When the process is working normally, only 2 per cent of accounts are mailed outside the specified period; that is, 2 per cent are 'defective'.

Control limits for the process can be calculated as follows:

$$\text{Mean proportion defective, } \bar{p} = 0.02$$

$$\text{Sample size } n = 1{,}000$$

$$\text{Standard deviation } s = \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

$$= \sqrt{\frac{0.02(0.98)}{1{,}000}}$$

$$= 0.0044$$

With the control limits at $\bar{p} \pm 3s$:

$$\text{Upper control limit (UCL)} = 0.02 + 3(0.0044) = 0.0332$$

$$= 3.32\%$$

$$\text{and lower control limit (LCL)} = 0.02 - 3(0.0044) = 0.0068$$

$$= 0.68\%$$

Figure 13.16 shows the company's control chart for this measure of quality over the last few weeks, together with the calculated control limits. It also shows that the process is in control.

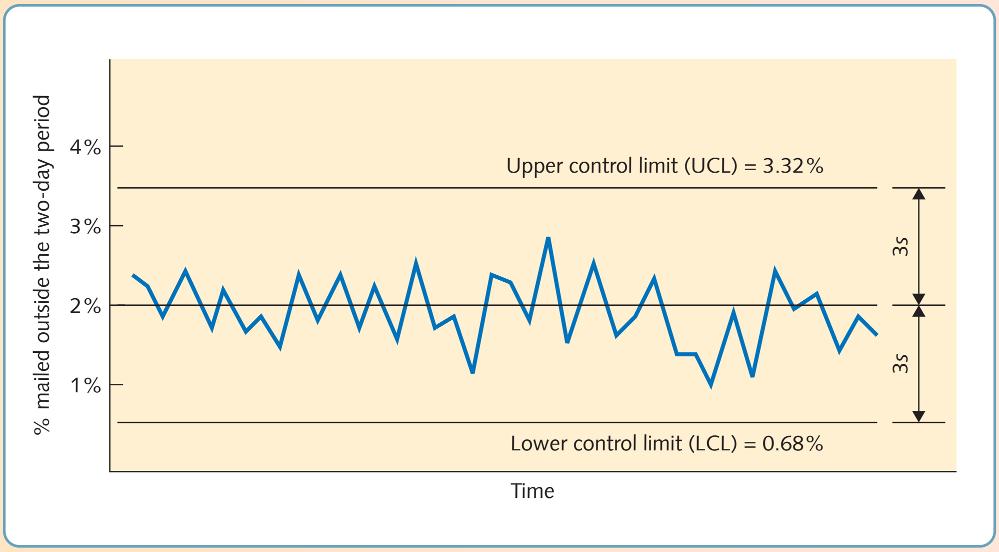


Figure 13.16 Control chart for the percentage of customer accounts that are mailed outside their two-day period

Sometimes it is more convenient to plot the actual number of defects (c), rather than the proportion (or percentage) of defectives, on what is known as a c-chart. This is very similar to the p-chart but the sample size must be constant and the process mean and control limits are calculated using the following formulae:

$$\text{Process mean } \bar{c} = \frac{c_1 + c_2 + c_3 \ldots c_m}{m}$$

$$\text{Control limits} = \bar{c} \pm 3\sqrt{c}$$

where c = number of defects
m = number of samples.

Applying the principles

1. An animal park in Amsterdam has decided to sample 50 visitors each day (n) to see how many visitors are from overseas. The data in Table 13.10 is for the last seven days. If it decided to continue recording this data and plot it on a control chart for attributes, what should the upper and lower control limits be?
2. The manager of a sweet shop decides to sample batches of sweets to check that the weight is reasonably consistent. She takes 9 samples, each with 10 bags. The data in Table 13.11 shows the average mean weight for each sample and the weight range. What control limits would a control chart for variables use?

Table 13.10 Number of visitors from overseas

Day	*Number of overseas visitors*
1	7
2	8
3	12
4	5
5	5
6	4
7	8

Table 13.11 The average mean weight for each sample and the weight range

Sample	*Weight average in grams $\bar{x}$*	*Range (R)*
1	10	1.50
2	8	2.00
3	9	3.00
4	9	2.50
5	8	1.50
6	9	1.00
7	11	2.00
8	14	2.50
9	12	2.00

Taking it further

Dale, B.G., Bamford, D. and van der Wiele, T. ***(2016) Managing Quality, 6th edition, Wiley.*** *This is the sixth edition of a book that has long been one of the best-respected texts in the area. A comprehensive and balanced guide.*

Oakland, J.S. and Oakland, R.J. ***(2018) Statistical Process Control, 7th edition, Routledge.*** *Another classic. Comprehensive, but not too daunting.*

Setter, C.J. ***(2018) Six Sigma: A complete step-by-step guide: A complete training & reference guide for white belts, yellow belts, green belts, and black belts, Harmony Living, LLC.*** *Very much the 'authorised' version from the Council for Six Sigma Certification*.

Webber, L. and Wallace, M. ***(2007) Quality Control for Dummies, John Wiley & Sons.*** *Written for the practitioner, this is an easy guide to improving quality. It covers expert techniques for introducing quality methods*.

Yadav, N. ***(2017) Implementation of Statistical Process Control (SPC) for Manufacturing, LAP LAMBERT Academic Publishing.*** *An up-to-date treatment, but relatively technical and manufacturing focused*.

14 Risk and resilience

Introduction

All operations face risks, experience external shocks and suffer internal failures. Some risks may emerge within the operation through, among other things, poor health-and-safety procedures, poorly designed quality processes, capacity constraints or short-term financial pressures. Some risks emerge as a consequence of supply network design – for example, reliance on over-powerful or unreliable suppliers, supply complexity, lack of visibility or uncertainty in demand. Other risks come from broader environmental forces, such as economic downturns, political unrest, pandemics and climate-related events. In the face of such risks, a 'resilient' operation or process is one that can identify likely sources of risks, prevent failures occurring or minimise their effects and learn how to recover from them. In a world where the pursuit of economic efficiencies can also mean greater vulnerabilities, managing risk is an increasingly vital task for operations managers (see Figure 14.1).

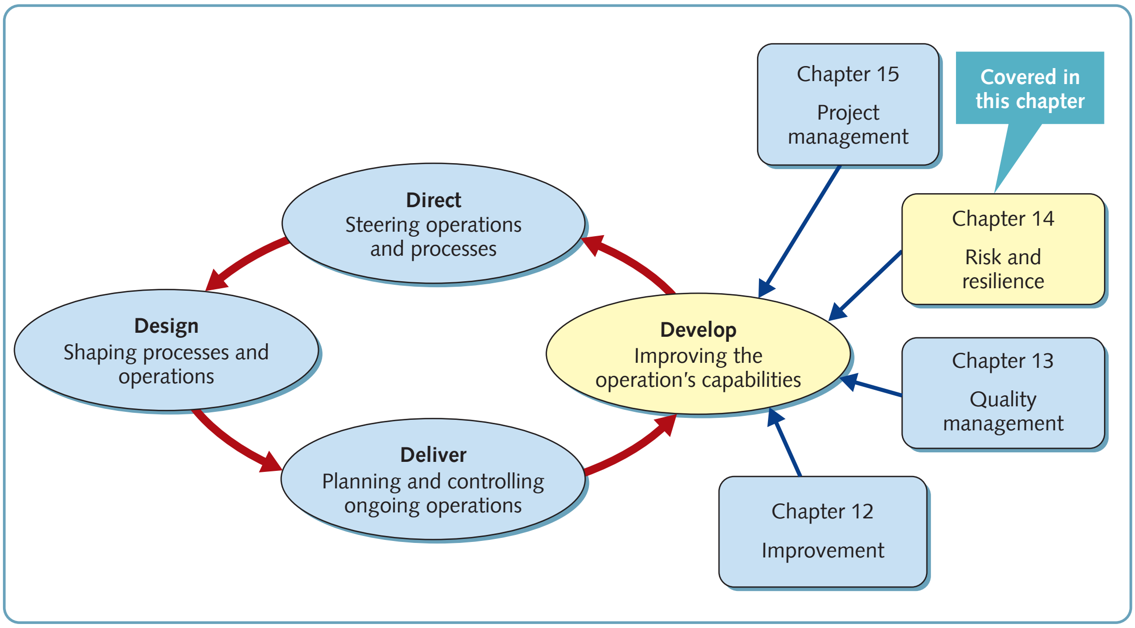


Figure 14.1 **Risk is the negative consequence of events; resilience is preventing, mitigating and recovering from these events**

EXECUTIVE SUMMARY

14.1 Is there an integrated approach to operations risk and resilience?

↓

14.2 Have potential risks been assessed?

↓

14.3 Have risk prevention measures been implemented?

↓

14.4 Have risk mitigation measures been implemented?

↓

14.5 Have risk recovery measures been implemented?

14.1 Is there an integrated approach to operations risk and resilience?

Risk is the potential for unwanted negative consequences from events. Resilience is the ability to prevent, withstand and recover from those events. Failure events can be categorised in terms of the seriousness of their impact and the likelihood of their occurrence. Relatively low-impact failures that happen relatively frequently are the province of quality management (see Chapter 13). Resilience involves attempting to reduce the combined effects of a failure occurring and the negative impact that it may have. However, there can be a trade-off between the efficiency of an operation and its resilience. Comprehensive risk-and-resilience management involves integrating four sets of activities: understanding and assessing the seriousness of the potential failures; preventing failures; minimising their negative consequences (called failure mitigation); and recovering from failure so as to reduce its impact.

14.2 Have potential risks been assessed?

Managing risk begins by identifying potential sources of risk. These sources can be categorised into supply failures, failures that happen inside the operation (further categorised as human, organisational and technology failures), cybersecurity, product/service design failures, customer failures and the failures caused by environmental forces such as politics, conflicts, weather, pandemics, and so on. Understanding why such failures occur can be aided by post-failure analysis using accident investigation, traceability, complaint analysis, fault-tree analysis and other similar techniques. Judging the likelihood of failure may be relatively straightforward for some well-understood processes, but often has to be carried out on a subjective basis, which is rarely straightforward.

14.3 Have risk prevention measures been implemented?

Once the potential sources of failure have been identified, managers must examine ways to prevent or reduce the likelihood of failures. The main approaches to failure prevention involve designing out the possibility of failure at key points in a process; installing fail-safeing mechanisms, which prevent the mistakes that can cause failure; maintaining processes so as to reduce the likelihood of failure; and providing extra but redundant resources that can provide back-up in the case of failure. Organisationally, these issues are often seen as part of occupational health and safety (OHS).

14.4 Have risk mitigation measures been implemented?

Even when a failure has occurred, its impact on the customer can, in many cases, be minimised through mitigation actions. Failure mitigation means isolating a failure from its negative consequences. There are various mitigation actions, including economic mitigation using insurance; risk sharing and hedging; spatial or temporal containment that prevents the failure spreading geographically or over time; loss reduction that removes whatever might be harmed by a failure; and substitution that involves providing substitute resources to work on a failure before it becomes serious.

14.5 Have risk recovery measures been implemented?

When failure occurs and cannot (or can only partially) be mitigated, operations managers must carefully consider alternative methods of recovery. Failure recovery is the set of actions that are taken to reduce the impact of failure once its negative effects have been experienced by the customer. Recovery needs to be planned and procedures put in place that can discover when failures have occurred, guide appropriate action to keep everyone informed, capture the lessons learned from the failure and plan to absorb these lessons into any future recovery.

14.1 Diagnostic question: *Is there an integrated approach to operations risk and resilience?*

Managing risk and resilience is one of any practising operations manager's most important roles. Risk is the potential for unwanted negative consequences from some event. Resilience is the ability to prevent, withstand and recover from those events. Failures (things that have negative consequences) happen in all operations, but accepting that failures occur is not the same thing as tolerating or ignoring them. Operations do generally attempt to minimise both the likelihood of failure and the effect it will have, although the method of coping with failure will depend on how serious are its negative consequences, and how likely it is to occur. This chapter is concerned with all types of failure other than those with relatively minor consequences. This is illustrated in Figure 14.2.

At the lower end of the scale, the whole area of quality management is concerned with identifying and reducing relatively small errors in the creation and delivery of products and services. This is not to minimise the importance of quality management. Continuing poor quality can still have a negative impact on any operation's reputation. Yet, other failures will have more immediate impact, even if they do not occur very frequently. For example, a server failure can seriously affect service and therefore customers, which is why system reliability is such an important measure of performance for IT service providers. Some failures are both relatively likely and have a significant impact. These failures are close to the top right-hand corner of the matrix in Figure 14.2. They are normally avoided by all businesses because embracing such risks would be clearly foolish. For example, building a factory in a war zone may be regarded as acceptable only if the pay offs from avoiding disaster outweigh the risks. Presumably such risks would be within the scope of a company's operations strategy, otherwise it would be 'betting the company' on the outcome. Finally, some failures, while much less likely, are serious in terms of their negative consequences. These are sometimes called 'black swan' events, a term that indicates an unpredictable event that is beyond what is normally expected yet has potentially severe consequences.[1] Examples may include major floods, earthquakes and hurricanes, wars and acts of terrorism, extreme financial collapses and pandemics.

OPERATIONS PRINCIPLE
Failure will always occur in operations; recognising this does not imply accepting or ignoring it.

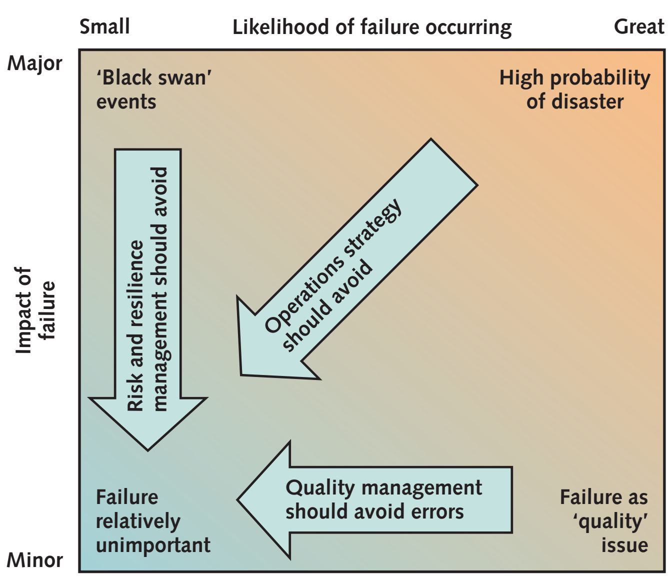


Figure 14.2 **How failure is managed depends on its likelihood of occurrence and the negative consequence of the failure**

Resilience versus efficiency

One of the most difficult issues to deal with in risk management is the trade-off between efficiency and resilience. During normal times, the objective of most operations is to meet their customers' requirements as efficiently as possible. The more predictable the conditions under which an operation works, the easier it is incrementally to pare away at surplus resources, such as surplus stocks, surplus operating capacity and the surplus administrative effort of managing multiple suppliers. Many of the previous chapters in this text have highlighted the efficiency gains to be obtained from such practices as 'optimising' inventory levels, matching capacity with (forecast) demand, outsourcing to low-cost regions, and so on. However, there can be a price to pay for such efficiencies. They can make operations susceptible to 'away-from-normal' disruptions. For example, most automobile manufacturers have set up elaborate and complex supply networks and honed their supply to all but eliminate inventory. The Honda plant at Swindon in the UK, for instance, requires 350 lorries full of critical components to cross the English Channel every day. The company told a UK Parliament committee that a delay of only 15 minutes costs it £850,000.[2] Such efficiency keeps prices low and tends to promote improvement, leading to further efficiency (see Chapter 11 on lean synchronisation), but it can also reduce an operation's ability to absorb shocks.

A similar argument applies to capacity. Like inventory, it can be seen as a redundant resource that needs reducing to make savings and hence enhance profits. Excess capacity can be represented as the enemy of efficiency. Yet, looked at another way, it gives the room for manoeuvre that will be needed if a sudden change in demand (or sometimes supply), or mix of activities, requires flexibility. During 'normal' times, when economic conditions are more or less 'as expected', most firms pursue efficiency over resilience. But after or during a severe disruption, the trade-off swings the other way. For example, during the COVID-19 pandemic, supply-chain disruption was mentioned nearly 30,000 times in the earnings calls (conversations with analysts, investors and the media to discuss a company's financial results) of the world's 2,000 largest listed firms, up from 23,000 in the same period the previous year. Mentions of 'efficiency' declined from 8,100 to 6,700.[3] During this period, the World Economic Forum called for companies to focus on 'risk competitiveness' just as they do on 'cost competitiveness'. However, companies often manage 'one shock at a time', and under 'normal' competitive pressures can return to their old ways of thinking about the efficiency–resilience trade-off.

OPERATIONS PRINCIPLE
There can be a trade-off between the resilience of an operation and its 'normal' efficiency.

Organisational slack

A concept from economics and organisational theory that is closely related to the idea of surplus resources providing a buffer to reduce risk, is that of 'organisational slack'. The term is not short of definitions, but is broadly taken to mean resources that are not necessary for the normal functioning of the organisation, or a surplus of facilities, equipment, supplies or personnel (all of which make up 'capacity') relative to current demands. The idea of slack is often used by organisational economists in a negative sense to signify systems that are running inefficiently. However, it has been hypothesised that it also has benefits. Much of the extensive research into organisational slack is beyond the scope of this text, but according to some authorities, organisational slack has four primary functions.[4] First, it can be used as an encouragement to retain sought-after employees' services. Second, it can help with problem resolution and process improvement. Third, and most relevant to our discussion here, organisational slack can be a buffering mechanism that can be used to adapt to sudden changes in the environment or periods of crisis. Finally, it can help long-term creative or strategic behaviour.

Assessment, prevention, mitigation and recovery

Obviously, some businesses function in a riskier environment than others. Those operations with a high likelihood of failure and/or serious consequences deriving from that failure will need to give it more attention. But operations and process resilience is relevant to all

Case example

Oil spills and pandemics: there is still an efficiency versus resilience trade-off[5]

Two disasters: 10 years apart, but both devastating.

BP and the Deepwater Horizon disaster

When an explosion and fire ripped through Transocean's Deepwater Horizon drilling rig, working under contract for BP oil company's Macondo well in the Gulf of Mexico, 11 men lost their lives, 17 were injured, and more than three million barrels of oil leaked into the ocean. A pulse of gas shot up, buckling the drill pipe. The emergency valve designed to cap the well in case of an accident, the 'blowout protector,' failed, and the gas reached the drill rig, triggering the fatal explosion and prompting the leak. The incident drew attention to the risks of drilling for oil in one of the most difficult, but ecologically rich, parts of the world. Drilling for oil in deep offshore waters, like the Gulf of Mexico, is inherently dangerous for the people working the platforms. But conditions on the Deepwater Horizon rig were particularly concerning. After the spill, an official US investigation that investigated the spill found that there were many lapses in safety that had contributed to the disaster. There had been an overarching 'failure of management', as well as poor decisions that could be traced back to a culture both within BP and the industry more broadly, which did not value safety enough. BP was found guilty of 'gross negligence' the US court ruled, and it found that several decisions leading to the disaster were 'primarily driven by a desire to save time and money, rather than ensuring that the well was secure'. In the years after the Deepwater Horizon crisis, BP also had to cut costs, in part to pay for its legal fees and for the clean-up bills that eventually exceeded $60 billion.

COVID-19 pandemic

In late December 2019, a pneumonia of unknown cause was detected in Wuhan, China. It was reported to the World Health Organization Country Office on 31 December. The outbreak was declared a Public Health Emergency of International Concern on 30 January 2020. On 11 February 2020, the WHO announced a name for the new coronavirus disease: COVID-19. Researchers believed that the coronavirus began in bats, then jumped to an intermediary species that passed it to people. The virus, which could survive for days on some surfaces, could be spread among humans by way of respiratory droplets within two metres or more. Its pneumonia-like symptoms include fever and coughing, causing death in some cases. Its economic symptoms were just as severe. Economic activity dropped drastically around the world. Its impact was variously described as the worst since the Second World War, the Great Depression of the 1920s and 30s, or the Black Death of 1346–1353 (take your pick). Yet it was undeniably serious for most businesses and governments. As well as hundreds of thousands of people dying, complex but fragile supply chains were disrupted; governments locked down whole industries to prevent the virus spreading; individual organisations closed their locations to preserve staff safety; millions had to work from home; and demand for most services and products shrank. It was this simultaneous impact on supply, operations processes and demand that caused such devastating economic impact, scarring economies for years afterwards.

How were they different?

The Deepwater Horizon disaster was largely a function of the internal actions (and mistakes) by the company and its contractors. The economic devastation spread as a result of the COVID-19 virus was an external event that came upon most organisations with little or no warning. Yet the negative outcomes of both events were partly a result of the efficiency–resilience trade-off – although in different ways. Clearly the cost pressures in the Deepwater Horizon well were excessive. No operation should allow its drive for efficiency to override reasonable safety concerns. Yet, no operations activity is entirely risk-free. The impact of COVID-19 could have been mitigated, but only for some businesses, and only to some extent. More robust supply arrangements, high in-process and finished goods stocks, and experience with flexible working arrangements could have helped some, but not all, operations. However, such strategies that pay off liberally in the event of even the worst case are usually prohibitively expensive.

Ken Cedeno/Corbis/Getty Images

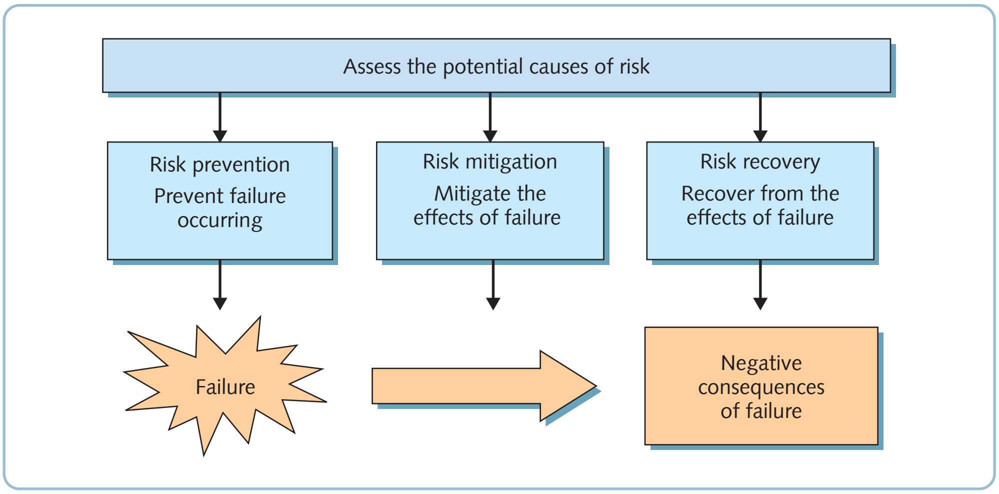


Figure 14.3 Managing operations and process risk involves assessing the potential causes of risk, failure prevention; mitigating the negative consequences of failure and failure recovery

organisations. They all must give attention to the four sets of activities that will determine their resilience. The first is concerned with understanding what failures could potentially occur in the operation and assessing their seriousness. The second is to examine ways of preventing failures occurring. The third is to minimise the negative consequences of failure (called 'mitigation'). The final task is to devise plans and procedures that will help the operation to recover from failures when they do occur. The remainder of this chapter deals with these four tasks (see Figure 14.3).

OPERATIONS PRINCIPLE
Risk is governed by the effectiveness of failure assessment, prevention, mitigation and recovery.

14.2 Diagnostic question: *Have potential risks been assessed?*

OPERATIONS PRINCIPLE
A 'failure to understand failure' is often the root cause of a lack of resilience.

A prerequisite to achieving operations and process resilience is to understand where risks might occur and what the consequences might be, by assessing causes of failure – often, it is the 'failure to understand failure' that leads to excessive disruption.

Identifying the potential sources of risk

The causes of some failure are purely random, like lightning strikes, and are difficult if not impossible to predict. However, the vast majority of failures are not like this – they are caused by something that could have been avoided, which is why, as a minimum starting point, a simple checklist of failure causes is useful. Figure 14.4 illustrates how this might be done. Here, failure sources are classified as: (1) failures of supply, internal failures such as those deriving from human organisational and technological sources; (2) failures deriving from the design of products and services; (3) failures deriving from customer failures; and (4) general environmental (or institutional) failures.

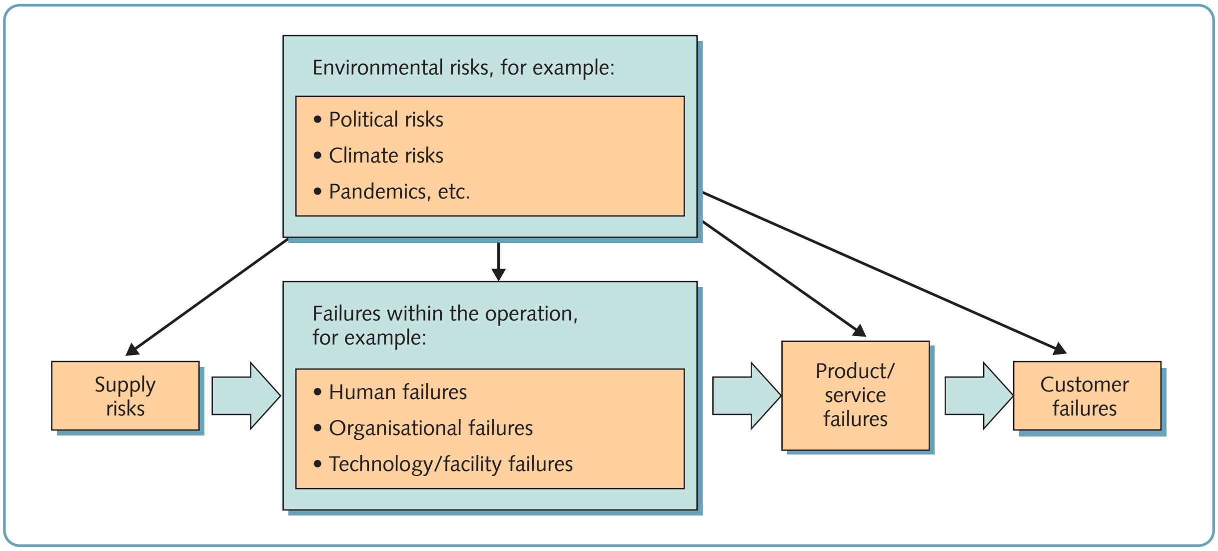


Figure 14.4 **The sources of potential risk for operations**

Supply risks

Supply failure means any failure in the timing or quality of products and services delivered into an operation – for example, suppliers delivering the wrong or faulty components, outsourced call centres suffering a telecoms failure, disruption to power supplies, and so on. The more an operation relies on suppliers for materials or services, the more it is at risk from failure caused by missing or substandard inputs. It is an important source of failure because of the increasing dependence on outsourced activities in many industries, and the emphasis on keeping supply chains 'lean' in order to cut costs. Other factors have also increased exposure to supply failure in recent years. For example, the rise of global sourcing usually means that parts are shipped around the world on their journey through the supply chain. Microchips manufactured in Taiwan could be assembled onto printed circuit boards in Shanghai that are then finally assembled into a computer in Ireland and delivered to the United States.

Human risks

There are two broad types of risks that are caused by human failure. The first is where key personnel leave, become ill, die, or in some way cannot fulfil their role. The second is where people are actively doing their job but are making mistakes. Understanding risk in the first type of failure involves identifying the key people without whom operations would struggle to operate effectively. These are not always the most senior individuals, but rather those fulfilling crucial roles that require special skills or tacit knowledge. Human failure through 'mistakes' also comes in two types: errors and violations. 'Errors' are mistakes in judgement: with hindsight, a person should have done something different, such as if the manager of a sports stadium fails to anticipate dangerous crowding. 'Violations' are acts that are clearly contrary to defined operating procedure, such as if a maintenance engineer fails to clean a filter in the prescribed manner. Catastrophic failures are often caused by a combination of errors and violations. For instance, one kind of accident, where an aircraft appears to be under control and yet still flies into the ground, is very rare (once in two million flights). For this type of failure to occur, first, the pilot has to be flying at the wrong altitude (error). Second, the co-pilot would have to fail to cross-check the altitude (violation). Third, air traffic controllers would have to miss the fact that the plane was at the wrong altitude (error). Finally, the pilot would have to ignore the ground proximity warning alarm in the aircraft, which can be prone to giving false alarms (violation).

Organisational risks

Organisational risks are usually taken to mean systemic failures that derive from a business's organisational leadership, structure and culture. This is a huge potential source of risk and includes almost all operations and process management. Examples include failure in the design of processes, such as bottlenecks causing system overloading, and failures in the resourcing of processes, such as insufficient capacity being provided at peak times. Ill-judged policy can also make failure more likely. For example, remuneration policy may motivate staff to work in a way that, although increasing the financial performance of the organisation, also increases risks. Sales people may be incentivised so that they make promises to customers that cannot be fulfilled. Investment bankers may be more concerned with profit (up-side risk) than the risks of financial overexposure (down-side risk). This type of risk can derive from an organisational culture that minimises consideration of risk, or it may come from a lack of clarity in reporting relationships.

Technology/facilities risks

By 'technology and facilities' we mean all the IT systems, machines, equipment and buildings of an operation. All are liable to failure, or breakdown. The failure may be only partial, such as a machine that has an intermittent fault. Alternatively, it can be what we normally regard as a 'breakdown' – a total and sudden cessation of operation. Either way, its effects could bring a large part of the operation to a halt. For example, a computer failure in a supermarket chain could paralyse several large stores until it is fixed. The root cause of these failures can often be traced back to inadequate maintenance of exceptionally complex information systems.

Cybersecurity risks

Any advance in processes or technology creates risks. No real advance comes without risk, threats and even danger. A specific type of technological failure is the failure of an operation's technology leading to exposure to cyber risk, which the Institute of Risk Management defines as 'any risk of financial loss, disruption or damage to the reputation of an organisation from some sort of failure of its information technology systems'. With the increased reliance on internet-based communication in all types of business, it has become a major risk factor. We will return to this type of risk later in the chapter.

Product/service design risks

Products and services often look great in design, but only when they have to cope with real demand do inadequacies become evident. An example is Heathrow's Terminal 5, which reported that it had not been able to replace a single light bulb in the five years since its opening, because in its design no one thought to examine how its 120,000 light fittings would be maintained when they reached the end of their life. After several failed attempts at finding a solution, Heathrow Airport Holdings announced they would be hiring the services of a specialist high-level rope-work company to carry out the work.[6] Of course, during the design process, potential risk of failure should have been identified and 'designed out'. But one only has to look at the number of 'product recalls' or service failures to understand that design failures are far from uncommon.

Customer risks

Not all failures are (directly) caused by the operation or its suppliers. Customers may 'fail' in that they misuse products and services. For example, an IT system might have been well designed, yet the user could treat it in a way that causes it to fail. Customers are not 'always right'; they can be inattentive and incompetent. However, merely complaining about customers is unlikely to reduce the chances of this type of failure occurring. Most organisations will accept that they have a responsibility to educate and train customers, and to design their products and services so as to minimise the chances of failure.

Case example

Volkswagen and the 'Dieselgate' scandal[7]

Some 'organisational failures' can border on criminality. What became known as 'Dieselgate' started out as a scandal that only affected Volkswagen, Germany's largest car company. But it eventually grew until it became a global issue involving many auto industry players. Some believe it even influenced governments' transport policies, legislation on air pollution and the move towards electric and hybrid cars. It started (or became evident) when the US Environmental Protection Agency (EPA) found that Volkswagen (VW) had been installing a piece of software into computers on its cars that falsified emissions data on its vehicles with diesel engines. The software (a so-called 'defeat device') could recognise when a car was being tested so that it could retune the engine's performance to limit nitrogen oxide emissions. After any test, when the car returned to normal road conditions, the level of such emissions increased sharply. The number of cars with the devices that enabled them to falsify their emission levels was estimated at 11 million cars worldwide. As the scale of the scandal became evident, VW's US boss did admit that they had 'totally screwed up', and the group's then chief executive said that VW had 'broken the trust of our customers and the public' (he later resigned because of the scandal). Over a year after the news broke, the US Department of Justice announced that Volkswagen was to pay \$4.3 billion under a plea deal with US authorities. This was in addition to a \$15 billion civil settlement with car owners and environmental authorities in the US, where VW had agreed to buy back some of the affected vehicles. However, appearing before a UK parliamentary committee, the head of Volkswagen UK said that the company had not misled anybody, nor did it owe any compensation to British drivers because it did not falsify information, nor mislead anybody. In May 2020, five years after the scandal broke, the German Federal Court of Justice ruled that Volkswagen car owners were entitled to damages because of the emissions scandal. It said that owners could return their car and receive the price paid minus a share for using the car in the intervening period. The ruling was thought likely to influence many other current court cases on the issue in Germany.

Environmental risks

Environmental disruption includes all the causes of failure that lie outside an operation's direct influence. Typically, such disasters include political upheaval, trade wars, weather events, fires, pandemics, corporate crime, terrorism and other security attacks. This source of potential failure has risen to near the top of many firms' agendas due to a series of major events over recent years. As operations become increasingly integrated (and increasingly dependent on integrated technologies such as information technologies), businesses are more aware of the critical events and malfunctions that have the potential to interrupt normal business activity and even stop the entire organisation.

Post-failure analysis

While sources of risk can often be identified in advance of their occurrence, it is also valuable to use previous failures to learn about sources of potential risk. This activity is called post-failure analysis, and could include the following:

- *Accident investigation* – Large-scale disasters like aeroplane accidents are usually investigated using accident investigation, where specifically trained staff analyse the causes of the accident. That so much attention goes into examining these kinds of failures is not only because of the damaging consequences of failure, but also because their infrequency makes it relatively hard to identify new sources of risk in advance of an event.
- *Failure traceability* – Some businesses (either by choice or because of a legal requirement) adopt traceability procedures to ensure that all their failures (such as contaminated food products) are traceable. Bio-tagging in pharmaceuticals is one example of this approach.

- *Complaint analysis* – Complaints (and compliments) are a potentially valuable source for detecting the sources of failures. The prime function of complaint analysis involves analysing the number and 'content' of complaints over time to understand better the nature of the failure, as the customer perceives it. However, for every customer who does complain, there may be many who do not.
- *Fault-tree analysis* – This is a logical procedure that starts with a failure or a potential failure and works backwards to identify all the possible causes and therefore the origins of that failure. Fault-tree analysis is made up of branches connected by two types of nodes: AND nodes and OR nodes. Figure 14.5 shows a simple fault tree identifying the possible reasons for a filter in a heating system not being replaced when it should have been.

Likelihood of risks occurring

The difficulty of estimating the chance of a risk occurring varies greatly. Some risks are the result of well-understood phenomena. A combination of rational causal analysis and historical performance data can lead to a relatively accurate estimate of failure occurring. For example, a mechanical component in an air conditioning unit installed within an airport departure lounge may fail within 18 and 24 months of its installation in 95 per cent of cases. Other types of failure are far more difficult to predict. The chances of a fire in a bank's outsourced call centre are (hopefully) low, but how low? There will be some data concerning fire hazards in this type of operation, and one may insist on regular hazard inspection reports from the supplier's insurance providers, but the estimated probability of failure will be both low and subjective. Assessing operational risks is not always easy. Even businesses that assess risks as part of their service can get it wrong. In what was one of the more ironic incidents, a worker at an insurance company reached an out-of-court compensation settlement after she fell over a pile of accident claim forms.

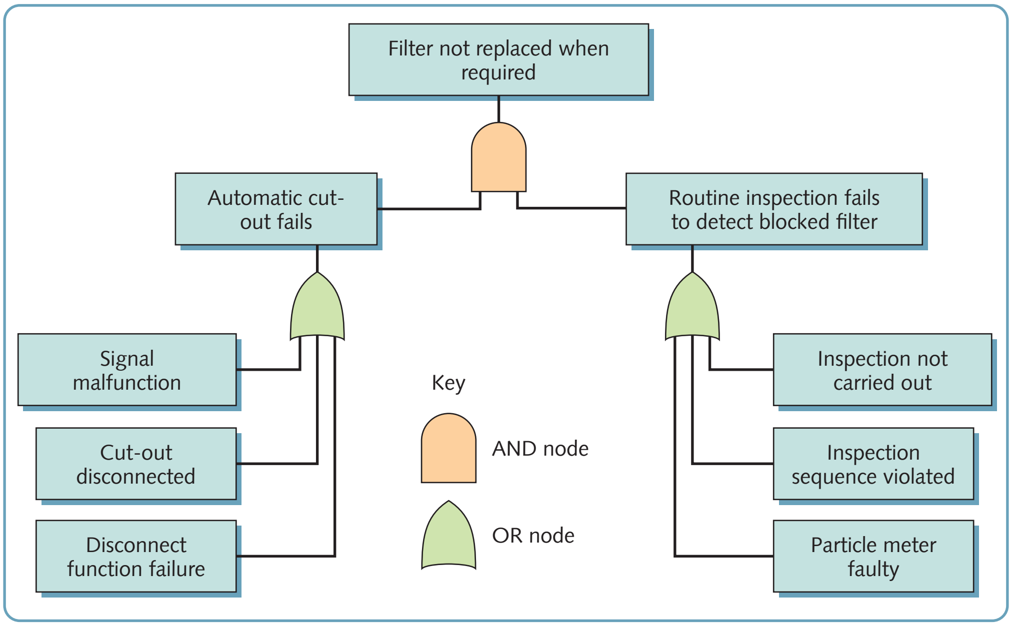


Figure 14.5 Fault-tree analysis for failure to replace filter when required

'Objective' estimates of risk

Estimates of failure based on historical performance can be measured in several ways, including:

- *failure rates* – how often a failure occurs;
- *reliability* – the chances of a failure occurring;
- *availability* – the amount of available useful operating time.

'Failure rate' and 'reliability' are different ways of measuring the same thing: the propensity of an operation, or part of an operation, to fail. Availability is one measure of the consequences of failure in the operation. Sometimes failure is a function of time. For example, the probability of an electric lamp failing is relatively high when it is first used, but if it survives this initial stage, it could still fail at any point and the longer it survives, the more likely its failure becomes. Most physical parts of an operation behave in a similar manner. The curve that describes failure probability of this type is called the bath-tub curve. It comprises three distinct stages:

1. The 'infant-mortality' or 'early-life' stage where early failures occur, caused by defective parts or improper use.
2. The 'normal-life' stage when the failure rate is usually low and reasonably constant, and caused by normal random factors.
3. The 'wear-out' stage when the failure rate increases as the part approaches the end of its working life, and failure is caused by the ageing and deterioration of parts.

Figure 14.6 illustrates three bath-tub curves with slightly different characteristics. Curve A shows a part of the operation that has a high initial infant-mortality failure but then a long, low-failure normal life followed by the gradually increasing likelihood of failure as it approaches wear-out. Curve B, while having the same stages, is far less predictable. The distinction between the three stages is less clear, with infant-mortality failure subsiding only slowly and a gradually increasing chance of wear-out failure. Failure of the type shown in curve B is far more difficult

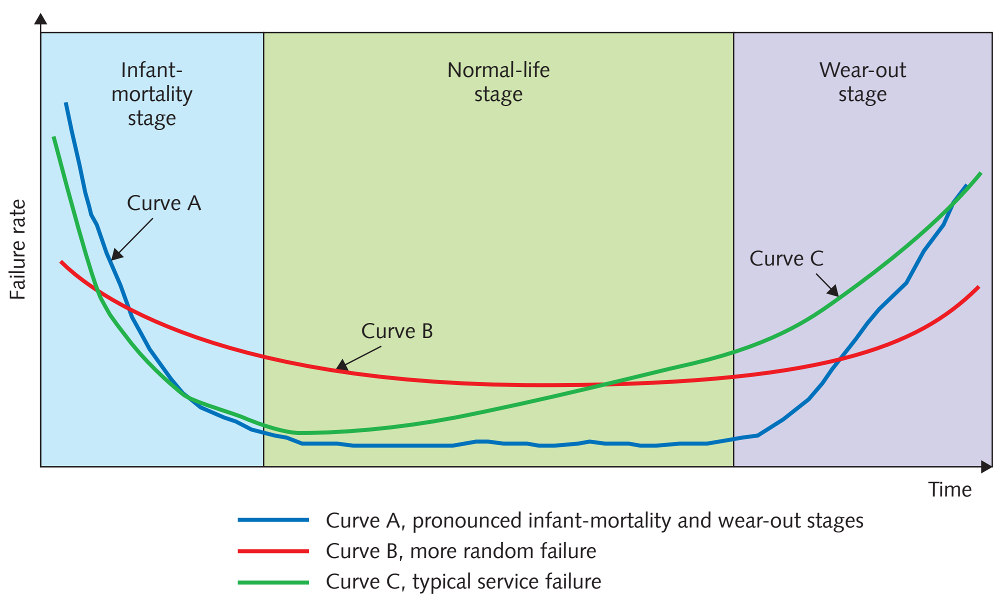


Figure 14.6 Bath-tub curves for three types of process

to manage in a planned manner. The failure of operations that rely more on human resources than on technology, such as some services, can be closer to curve C of Figure 14.6. They may be less susceptible to component wear-out but more so to staff complacency. Without review and regeneration, the service may become tedious and repetitive, and after an initial stage of failure reduction, as problems in the service are ironed out, there can be a long period of increasing failure.

'Subjective' estimates of risk

Failure assessment, even for subjective risks, is increasingly a formal exercise that is carried out using standard frameworks, often prompted by health-and-safety concerns, environmental regulations, and so on. However, individual attitudes to risk are complex and subject to a wide variety of influences. In fact, many studies have demonstrated that people are generally very poor at making risk-related judgements. Consider the success of state and national lotteries: the chances of winning, in nearly every case, are extraordinarily low, and the costs of playing sufficiently significant to make the financial value of the investment entirely negative. If a player has to drive their car in order to purchase a ticket, they may be more likely to be killed or seriously injured en route than they are to win the top prize. But, although people do not always make rational decisions concerning the chances of failure, this does not mean abandoning the attempt. However, it does mean that one must understand the limits of overly rational approaches to failure estimation.

> **OPERATIONS PRINCIPLE**
> ***Subjective estimates of failure probability are better than no estimates at all.***

Non-evident failure estimation

Not all failures are immediately evident. Small failures may accumulate for a while before they become evident, making objective and subjective estimation challenging. For example, purchasing managers who encounter difficulties in using an e-procurement system may simply circumvent the system, with the failure points only becoming evident when levels of non-compliance reach sufficient levels for senior management to notice. Likewise, within an automated production line, debris can accumulate. This may not cause an immediate failure, but could eventually lead to sudden and dramatic failure.

'Tech debt'

An example of failures only becoming evident a long time, sometimes decades, after they originated is what is sometimes called 'tech debt'. System outages, corrupt data, installation delays and poor service resulting from old and degraded software and faulty source codes can be a significant cause of risk. Source code errors represent a hidden risk to many businesses as IT becomes more central to operations processes. Tech debt (also known as technical debt, design debt or code debt) is what occurs when development teams expedite the delivery of some aspect of a software project that later needs to be repaired or patched. Although originally intended to explain the consequences of prioritising speedy delivery over perfect code, the term has come to be used to describe almost any coding or architecture error that 'lurks' in software, to become evident later. The concept has been used to highlight the difference between established retail banks, with many years of legacy code contained within their systems, and the newer 'challenger' banks that have started in some markets, which have newer, often better-integrated systems with fewer hidden 'legacy risks'.

Health and safety

One of the paradoxes of operations management is that occupational health and safety (OHS) management is such a large part of many practitioners' work, yet has attracted relatively little academic attention in the OM field. Most serious academic studies of the topic come from the human resources or industrial relations areas, with a growing interest from academic lawyers as legislation in the area has increased in most parts of the world. Not only is it a subject of significant practitioner interest, it is clearly important, both for the potential improvement of operations performance that excellent OHS practices can bring, and for its clear ethical benefits. Occupational accidents and diseases have very significant adverse consequences. Staff are injured, absent from work and possibly retired early, facilities are damaged, the quantity and quality of output declines, all of which has a negative impact on an operation's performance and reputation. It has been estimated that, globally, this type of incident results in nearly 2.3 million deaths every year and incurs over $2.8 trillion of costs.[8]

In many countries, there has been a growing body of acceptance that the adoption of systematic OHS management systems is necessary to ensure safe and productive work environments. Such systems often stress the importance of adopting the systematic identification of potential hazards, the assessment and control of risks, and evaluation and periodic review of risk measures, in fact many of the issues covered in this chapter. Operations managers have been known to complain about the amount of time and bureaucracy involved in OHS. However, health and safety management does not have to be complicated, costly or time-consuming. In fact, much of what is usually required by legislation is good risk management practice, including such issues as:

OPERATIONS PRINCIPLE
Ensuring the health and safety of all concerned with their operation's activities is a fundamental responsibility for operations managers.

- Identifying responsible individual(s).
- Formally recording OHS procedures.
- Identifying and controlling risks.
- Consulting staff.
- Providing adequate training.
- Providing appropriate equipment and facilities.
- Making arrangements for first aid, accidents and ill health.
- Investing in appropriate insurance (and other forms of risk mitigation).

Failure mode and effect analysis

Having identified potential sources of failure (either in advance of an event or through post-failure analysis) and having then examined the likelihood of these failures occurring through some combination of objective and subjective analysis, managers can move to assigning relative priorities to risk. The most well-known approach for doing this is failure mode and effect analysis (FMEA). Its objective is to identify the factors that are critical to various types of failure as a means of identifying failures before they happen. It does this by providing a 'checklist' procedure built around three key questions for each possible cause of failure:

1. What is the likelihood that failure will occur?
2. What would the consequence of the failure be?
3. How likely is such a failure to be detected before it affects the customer?

Based on a quantitative evaluation of these three questions, a risk priority number (RPN) is calculated for each potential cause of failure. Corrective actions, aimed at preventing failure, are then applied to those causes whose RPN indicates that they warrant priority (see Figure 14.7).

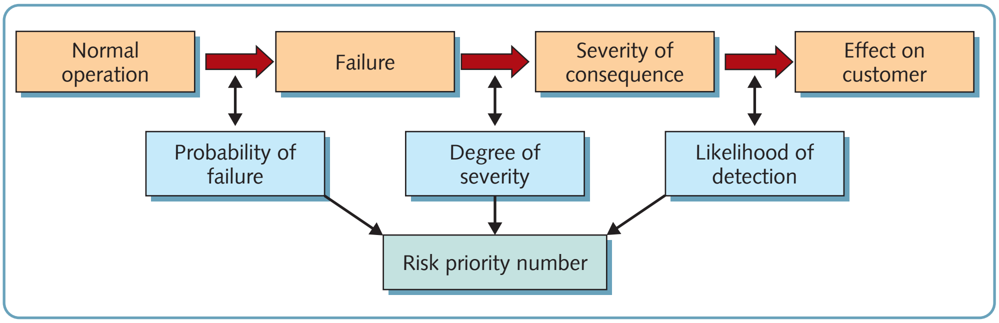


Figure 14.7 Procedure for failure mode and effect analysis (FMEA)

14.3 Diagnostic question: *Have risk prevention measures been implemented?*

It has been said many times, 'prevention is better than cure'. This is why risk prevention is such an important part of operations and process resilience. There are a number of approaches to this, including designing out failure points, fail-safeing, maintenance and deploying redundant resources.

Designing out fail points

Process mapping, described in Chapter 6, can be used to 'engineer out' the potential fail points in various operations. By examining the stages in the process map, and identifying the parts of the process that are particularly prone to failure and the stages that are critical to the success of the service, one can focus on potential critical issues. This is best done by the staff of the operation who understand the differences between the 'formal' process and what actually happens in practice. Metaphorically 'walking themselves through' the process and discussing each stage in turn helps to focus on key risks.

Case example

Thousands of customers locked out of their accounts in IT meltdown[9]

When Paul Pester, the CEO of British retail bank TSB (owned by Spain's Banco Sabadell), stepped down, he said he wanted the bank to be led 'without distraction'. He was taking responsibility for an IT meltdown the previous year that had cost the bank £176.4 million, pushed it into a half-year loss of over half a million pounds and was probably the most expensive IT problem in British banking history. The bank's problems started when five million customer accounts were migrated from its old IT system, inherited from the bank's previous owner, to one operated by Sabadell. The move had been planned for years and customers had been informed that TSB would shut down some of its services for two days so it could develop its new IT system, which was designed, it said, 'for the digital age'. The new system was expected to make savings of up to £100 million a year.

Unfortunately, the planned migration caused many of the bank's customers to be unable to make transactions or to see their balances. Almost two million customers lost access to their online banking services. Worse, the problems continued for several weeks. More than 130,000 complaints needed a team of more than 260 staff to deal with them. As it was working to try and put things right, TSB promised that no customer would lose money because of the problems. Several reasons were put forward for the cause of the disaster. Some blamed a lack of planning and especially a lack of sufficient testing of the new system; some claimed that the bank had greatly underestimated the complexity and scale of the project. Others said that Sabadell's internal IT service did not fully understand what was involved in the migration, which should have been phased in rather than changed in one 'big bang'. Whatever the cause (or more likely, causes), two things are certain: first, with hindsight, TSB was too optimistic about its ability to carry out such a complex and sensitive project; and second, such public IT failures can have a profound effect on an organisation's finances and reputation.

Fail-safeing

The concept of fail-safeing has emerged since the introduction of Japanese methods of operations improvement. Called *poka yoke* in Japan (from *yokeru* (to prevent) and *poka* (inadvertent errors)), the idea is based on the principle that human mistakes are to some extent inevitable. What is important is to prevent them from becoming defects. *Poka yokes* are simple (preferably inexpensive) devices or systems that are incorporated into a process to prevent inadvertent operator mistakes resulting in a defect. Typical *poka yokes* include:

> OPERATIONS PRINCIPLE
> ***Simple methods of fail-safeing can often be the most cost effective.***

- Trays used in hospitals with indentations shaped to each item needed for a surgical procedure – any item not back in place at the end of the procedure might have been left in the patient.
- Checklists that have to be filled in, either in preparation for or on completion of an activity, such as a maintenance checklist for a plane during turnaround.
- Gauges placed on machines through which a part has to pass in order to be loaded onto, or taken off, the machine – an incorrect size or orientation stops the process.
- The locks on aircraft lavatory doors, which must be turned to switch the light on.
- Beepers on ATMs to ensure that customers remove their cards, or in cars to remind drivers to take their keys with them.
- Limit switches on machines that allow the machine to operate only if a part is positioned correctly.
- Height bars on amusement rides to ensure that customers exceed size limitations.

Maintenance

While managers can try to design out failures and use fail-safe (*poka yoke*) mechanisms to further reduce the likelihood of failures, all operations and processes also need maintaining. Maintenance is the term typically used to cover the way operations and processes try to avoid failure by taking care of their physical facilities. It is particularly important when physical facilities play a central role in the operation, such as power stations, airlines and petrochemical refineries. There are a number of approaches to maintenance, including preventive maintenance, condition-based maintenance and total productive maintenance.

Preventive maintenance (PM)

This attempts to eliminate or reduce the chances of failure by regularly servicing facilities. For example, the engines of passenger aircraft are checked, cleaned and calibrated according to a regular schedule after a set number of flying hours.

Condition-based maintenance (CBM)

This attempts to perform maintenance only when the facilities require it. For example, continuous-process equipment, such as that used in coating photographic paper, is run for long periods in order to achieve the high utilisation necessary for cost-effective production. Here CBM might involve continuously monitoring the vibrations or some other characteristic of the line, and then using the data to decide when to replace the bearings.

Total productive maintenance (TPM)

This combines aspects of PM and CBM approaches, but has a very strong focus on allowing people to take more responsibility for maintenance tasks. The approach looks to create a clear plan for all maintenance activities, including the level of preventive maintenance that is required for each piece of equipment, the standards for CBM and the respective responsibilities of operating staff and maintenance staff. For this to work, all staff must be trained in relevant maintenance skills and have all the skills needed to carry out their roles. Finally, TPM looks to use information from maintenance activities to steadily design out maintenance by considering which maintenance activities are a consequence of poor design, manufacturing or installation.

How much maintenance?

Most operations plan their maintenance to include a level of regular preventive maintenance that gives a reasonably low but finite chance of breakdown. Usually, the more frequent the preventive maintenance episodes, the less are the chances of a breakdown. Infrequent preventive maintenance will cost little to provide but will result in a high likelihood (and therefore cost) of breakdown. Conversely, very frequent preventive maintenance will be expensive to provide but will reduce the cost of having to provide breakdown maintenance, as shown in Figure 14.8(a). The total cost of maintenance appears to minimise at an 'optimum' level of preventive maintenance. However, this may not reflect reality. The cost of providing preventive maintenance in Figure 14.8(a) assumes that it is carried out by a separate set of people (skilled maintenance staff) whose time is scheduled and accounted for separately from the 'operators' of the

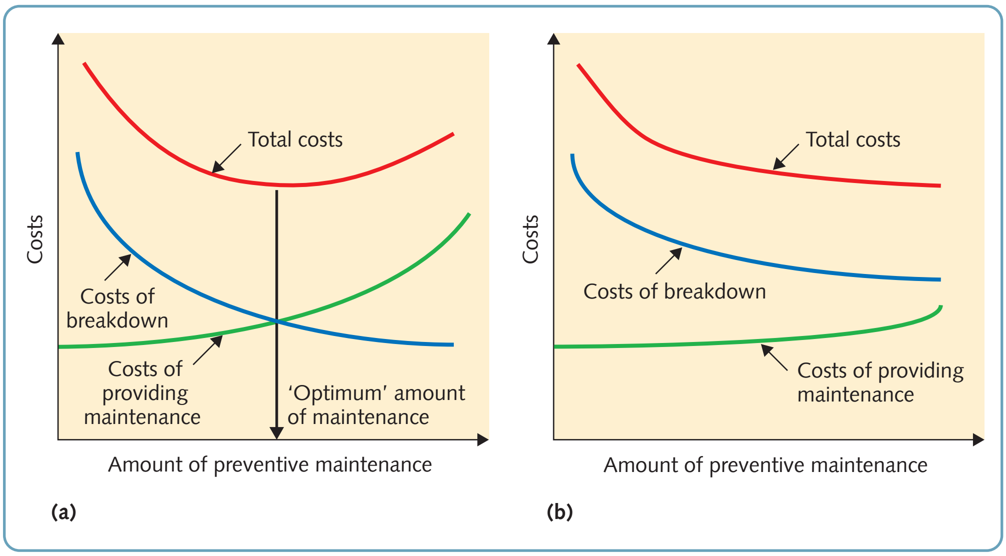


Figure 14.8 Two views of maintenance costs: (a) one model of the costs associated with preventive maintenance shows an optimum level of maintenance effort; and (b) if routine preventive maintenance tasks are carried out by operators and if the real cost of breakdowns is considered, the 'optimum' level of preventive maintenance shifts towards higher levels

facilities. In many operations, however, at least some preventive maintenance can be performed by the operators themselves (which reduces the cost of providing it) and at times that are convenient for the operation (which minimises the disruption to the operation). Furthermore, the cost of breakdowns could also be higher than is indicated in Figure 14.8(a) because unplanned downtime can take away stability from the operation, preventing it being able to improve itself. Put these two ideas together and the minimising total curve and maintenance cost curve look more like Figure 14.8(b). The emphasis is shifted towards using more preventive maintenance than is generally thought appropriate.

OPERATIONS PRINCIPLE
Preventive maintenance, by reducing unplanned downtime, can reduce the total costs of running equipment.

Redundancy

Building-in redundancy to an operation means having back-up processes or resources in case of failure. It can be an expensive solution to reduce the likelihood of the failure and is generally used when the breakdown could have a critical impact. Redundancy means doubling or even tripling some of the elements in a process so that these 'redundant' elements can come into action when a process fails. Nuclear power stations, hospitals and other public buildings have auxiliary or back-up electricity generators ready to operate in case the main electricity supply should fail. Some organisations also have 'back-up' staff held in reserve in case someone does not turn up for work, or is held up on one job and is unable to move on to the next. Spacecraft have several back-up computers on board that will not only monitor the main computer, but also act as a back-up in case of failure. Human bodies contain two of some organs – kidneys and eyes, for example – both of which are used in 'normal operation' but the body can cope with a failure in one of them. One response to the threat of large failures, such as terrorist activity, has been a rise in the number of companies (known as 'business continuity' providers) offering 'replacement office' operations, fully equipped with normal internet and telephone communications links, and often with access to a company's current management information. Should a customer's main operation be affected by a disaster, business can continue in the replacement facility within days or even hours.

The effect of redundancy can be calculated by the sum of the reliability of the original process component and the likelihood that the back-up component will both be needed and be working:

$$R_{a+b} = R_a + (R_b \times P_{(\text{failure})})$$

where

R_{a+b} = reliability of component a with its back-up component b
R_a = reliability of a alone
R_b = reliability of back-up component b
$P_{(\text{failure})}$ = the probability that component a will fail and therefore component b will be needed.

So, for example, an e-auction service provider has two servers, one of which will come online only if the first server fails. If each server has a reliability of 0.9, the two working together (each with reliability = 0.9) will have a reliability of:

$$0.9 + [0.9 \times (1 - 0.9)] = 0.99$$

Cybersecurity risks

Earlier in this chapter, we identified cybersecurity as a relatively new but increasingly serious risk. The internet is, by design, an open non-secure medium. Since the original purpose of the internet was not for commercial purposes, it is not designed to handle secure transactions. There is a trade-off between providing wider access through the internet, and the security concerns it generates. Three developments have amplified cybersecurity concerns. First, increased connectivity means that everyone has at least the potential to 'see' everyone else. Organisations want to make enterprise systems and information more available to internal employees, business partners and customers (see Chapter 10 on planning and control systems). Second, there is reduced 'perimeter' security as more people work from home or through mobile communications. Hackers hope to exploit lower levels of security in home computers in order to burrow into corporate networks. Third, it takes time to discover all possible sources of risk, especially as new technologies are introduced.

Cyber risk includes hacker attacks, data breaches, virus transmission, cyber extortion and network downtime. Any of these could be the result of:

- Deliberate unauthorised breaches of security firewalls.
- Unintentional/accidental breaches of security (which could still constitute a risk).
- Poor systems integrity in the design of IT systems.
- Unsecure use of cloud usage or storage.
- Unpoliced use of 'bring your own device' (BYOD) policies.

Most authorities on cybersecurity stress that, stripped of its technological terminology, cyber risk is just another sort of risk, which should be treated using the same identify, prevent, mitigate, recover framework that we are using in this chapter. The UK Government security services maintain that about 80 per cent of cyber attacks would be defeated by using these basic security controls. But, even so, there is still room for different approaches. While some experts stress the importance of 'perimeter' security, through the use of firewalls, others think perimeter defence insufficient when it comes to cyber threats (see the case example on Darktrace). As one commentator put it, 'If you put up a wall, I'm just going to show up with a ladder. If you put up barbed wire, I'm going to dig a hole under the wall. And, if all else fails, I'm just going to bribe somebody on the inside to open the door for me.'[10]

Case example

Darktrace uses AI to guard against cyber attacks[11]

As cyber criminals deploy increasingly sophisticated forms of attack to breach companies' defences and spread malicious, data-stealing codes, so the industry selling (hopefully) equally sophisticated protection services has grown. One of the most prominent among these is Darktrace, based in Cambridge in the UK, but with offices around the world. It was founded by mathematicians from the University of Cambridge and government cyber intelligence experts in the UK and the US. The company is well established in the application of artificial intelligence (AI) for cyber defence. One great advantage of AI is that it offers the possibility of keeping pace with the ever-evolving nature of cyber threats. Powered by what it describes as unsupervised machine learning, the company's AI responds to cyber threats before they become a crisis. Its self-learning technology is modelled on the human immune system, learning 'on the job', from the data and activity that it observes. AI and machine learning could potentially have a number of advantages over purely human surveillance. A skilled human could look for suspicious patterns, analyse them, devise ways to mitigate the threat, then inform the rest of the business, but this takes time. An analyst might need to spend anything between half an hour and half a day investigating one single suspicious security incident. And with the threat of cyber attacks rising there is a shortage of cybersecurity experts.

Darktrace says that its AI cybersecurity solution accelerates this process, conducting continuous investigations in the background to normal operating, at a pace and scale beyond the capabilities of a human analyst. Not only that, but AI-driven cybersecurity can handle specialist investigations into a large number of parallel threads

▶

concurrently, and instantly communicate its findings. However, some cybersecurity commentators believe that although AI offers the possibility of vastly improved security systems, like any innovative software development its effectiveness can sometimes be overstated. One problem with such systems is their tendency to show 'false positives – that is reporting a possible cybersecurity breach when there is actually no threat. Yet, even experts who suggest that there is excessive hype around AI admit that there is certainly a role for AI in cybersecurity, if only because it is particularly good at dealing with vast amounts of information and understanding what is normal and what is anomalous.

Michael S. Williamson/The Washington Post/Getty Images

14.4 Diagnostic question: *Have risk mitigation measures been implemented?*

Failure mitigation means isolating a failure from its negative consequences. It is an admission that not all failures can be avoided. However, in some areas of operations management, relying on mitigation rather than prevention is unfashionable. For example, 'inspection' practices in quality management were based on the assumption that failures were inevitable and needed to be detected before they could cause harm. Modern total quality management (TQM) places much more emphasis on prevention. Yet, in operations and process resilience, mitigation can be vital when used in conjunction with prevention in reducing overall risk.

Failure mitigation actions

The nature of the action taken to mitigate failure will obviously depend on the nature of the failure. In most industries, technical experts have established a classification of failure mitigation actions that are appropriate for the types of risk likely to be suffered. So, for example, in agriculture, government agencies and industry bodies have published mitigation strategies for such 'failures' as the outbreak of crop disease, contagious animal infections, and so on. Likewise, governments have different contingency plans in place to deal with the spread of major health risks. Although these types of plans tend to be industry specific, the following generic categorisation gives a flavour of the types of mitigation actions that may be generally applicable.

Mitigation planning – This is the activity of ensuring that all possible failure circumstances have been identified and the appropriate mitigation actions identified. It may be described in the form of a decision tree or guide rules. Almost certainly there will be some form of escalation that will guide the extra mitigation effort, should early actions not prove successful. Mitigation planning, as well an overarching action, also provides mitigation action in its own right. For example, if mitigation planning has identified appropriate training, job design, emergency procedures, and so on, then the financial liability of a business for any losses should a failure occur will be reduced. Certainly, businesses that have not planned adequately for failures will be more liable in law for any subsequent losses.

Economic mitigation – This includes actions to prevent losses from failure, spreading the financial consequences of failure and 'hedging' against failure. Insurance is the best known of these actions and is widely adopted, although ensuring appropriate insurance and effective claims management is a specialised skill in itself. Spreading the financial consequences of failure could involve, for example, spreading the equity holding in supply companies to reduce the financial consequences of such companies failing. Hedging involves creating a portfolio of ventures whose outcomes happen to be correlated so as to reduce total variability.

Containment (spatial) – This means stopping the failure from physically spreading to affect other parts of an internal or external supply network. Preventing contaminated food from spreading through the supply chain, for example, will depend on real-time information systems that provide traceability data.

Containment (temporal) – This means containing the spread of a failure over time. It particularly applies when information about a failure or potential failure needs to be transmitted without undue delay. For example, systems that give advance warning of hazardous weather such as snow storms must transmit the information to local agencies such as the police and road-clearing organisations, in time for them to stop the problem causing excessive disruption.

Loss reduction – This covers any action that reduces the catastrophic consequences of failure by removing the resources that are likely to suffer those consequences. For example, the road signs that indicate evacuation routes in the event of severe weather, or the fire drills that train employees in how to escape in the event of an emergency, may not reduce the consequences of failure on buildings or physical facilities, but can help dramatically in reducing loss of life or injury.

Substitution – This means compensating for failure by providing other resources that can substitute for those rendered less effective by the failure. It is a little like the concept of redundancy that was described earlier, but does not always imply excess resources if a failure has not occurred. For example, in a construction project, the risk of encountering unexpected geological problems may be mitigated by the existence of a separate work plan, which is invoked only if such problems are found. The resources may come from other parts of the construction project, which will in turn have plans to compensate for their loss.

> OPERATIONS PRINCIPLE
> ***If failure occurs, well-designed mitigation can reduce its negative impact on an operation and its customers.***

Table 14.1 gives some examples of each type of failure mitigation actions for three failures: the theft of money from one of a company's bank accounts; the failure of a new product technology to work adequately during the new product development process; and the outbreak of fire at a business premises.

Table 14.1 Failure mitigation actions for three failures

	Type of failure		
Failure mitigation actions	***Financial failure – theft from company account***	***Development failure – new technology does not work***	***Emergency failure – fire at premises***
Mitigation planning	Identify different types of theft that have been reported and devise mitigation actions, including software to identify anomalous account behaviour	Identify possible types of technology failure and identify contingency technologies, together with plans for accessing contingency technologies	Identify fire hazards and methods of detecting, limiting and extinguishing fires
Economic mitigation	Insure against theft and possibly use several different accounts	Invest in, or form partnership with, supplier of alternative technology	Insure against fire and have more, smaller, premises

Table 14.1 (*Continued*)

Failure mitigation actions	*Type of failure* Financial failure – theft from company account	Development failure – new technology does not work	Emergency failure – fire at premises
Containment (spatial)	'Ring-fence' accounts so a deficit in one account cannot be made good from another account	Develop alternative technological solutions for different parts of the development project so that failure in one part does not affect the whole project	Install localised sprinkler systems and fire door barriers
Containment (temporal)	Invest in software that detects signs of possible unusual account behaviour	Build-in project milestones that indicate the possibility of eventual development failure	Install alarm systems that indicate the occurrence of fire to everyone who may be affected (including in other premises)
Loss reduction	Build-in transfer delays until approval for major withdrawals has been given, also institute plans for recovering stolen money	Ensure the development project can use old technology if new one does not work	Ensure means of exit and employee training are adequate
Substitution	Ensure that reserve funds and staff to manage the transfer can be speedily brought into play	Have fall-back work package for devoting extra resources to overcome the new technology failure	Ensure back-up team that can take over from premises rendered inoperative by fire

Case example

Pressing the passenger panic button[12]

It is every nervous air traveller's nightmare – what if the pilot (or both pilots if there are two) becomes incapacitated? Such fear is not unwarranted. It is a real, although fortunately rare, danger with light aircraft. One Australian report identified 15 cases of the pilots of small aircraft being incapacitated in a five-year period. If there is an instructor on the ground to give an instant flying lesson to someone on board a plane, the plane and its temporary pilot may be talked down successfully. But it is a very difficult task, and can all too easily result in what is known euphemistically in aviation circles as a 'collision with terrain'. Keeping an aircraft flying straight and level while finding an appropriate bearing to arrive at a suitable airfield, and then landing safely takes a miracle. Which is why Garmin, the American multinational technology company, best known for its satellite-based navigation systems, has, in effect, developed a 'panic button' for passengers who find themselves in such a dangerous position. The company, which also makes electronic control systems for aircraft, developed Autoland, a panic button that switches control of the plane to its flight computers, in a similar way to engaging the autopilots used in commercial aircraft. But, in addition, Garmin's system also transmits an emergency radio code alert to air-traffic control and other planes in the area. It will also analyse other factors, such as weather conditions and the amount of available fuel before selecting an appropriate airfield to divert to. When the plane arrives at that airport, the system controls the descent and lands just as a human pilot would. Once on the ground, it even automatically applies the brakes, bringing the plane to a halt, and turns the engine off. During what must be a terrifying experience, passengers are kept informed about what is happening by way both of messages on a screen and voice announcements. They are also warned not to touch the controls, but to sit back and fasten their seat belts.

14.5 Diagnostic question: *Have risk recovery measures been implemented?*

Not all failure can be prevented or mitigated. So operations need to plan carefully how they might recover if failure does occur and when negative impacts have been experienced. Yet even where the customer sees a failure, it may not necessarily lead to dissatisfaction; customers may even accept that things occasionally do go wrong. If there is a metre of snow on the train lines, or if the restaurant is particularly popular, we may accept that the product or service does not work. It is not necessarily the failure itself that leads to dissatisfaction but often the organisation's response to the failure. Mistakes may be inevitable, but dissatisfied customers are not. For example, when a flight is delayed by five hours, there is considerable potential for dissatisfaction. But if the airline informs passengers that the aircraft has been delayed by a cyclone at its previous destination and that arrangements have been made for accommodation at a local hotel with a complimentary meal, passengers might then feel that they have been well treated and even recommend that airline to others. As such, a good recovery can turn frustrated customers into loyal ones.

The complaint value chain

The complaint value chain helps us to visualise the potential value of good recovery at different stages (see Figure 14.9). In Figure 14.9(a) an operation provides service to 5,000 customers, but 20 per cent experience some form of failure. Of these 1,000 customers, 40 per cent decide not to complain, perhaps because it seems like more trouble than it's worth or because the complaint processes are too convoluted. Evidence suggests that around 80 per cent of these non-complainers will switch to an alternative service provider (of course the precise switching percentage will depend on the number of alternatives in the market and the ease of switching). Another group of the 1,000 customers who experienced a failure do decide to complain, in this case 60 per cent. Some will be satisfied (in this case, 75 per cent) and others will not (in this case, 25 per cent). Dissatisfied complainers will generally leave the organisation (for this

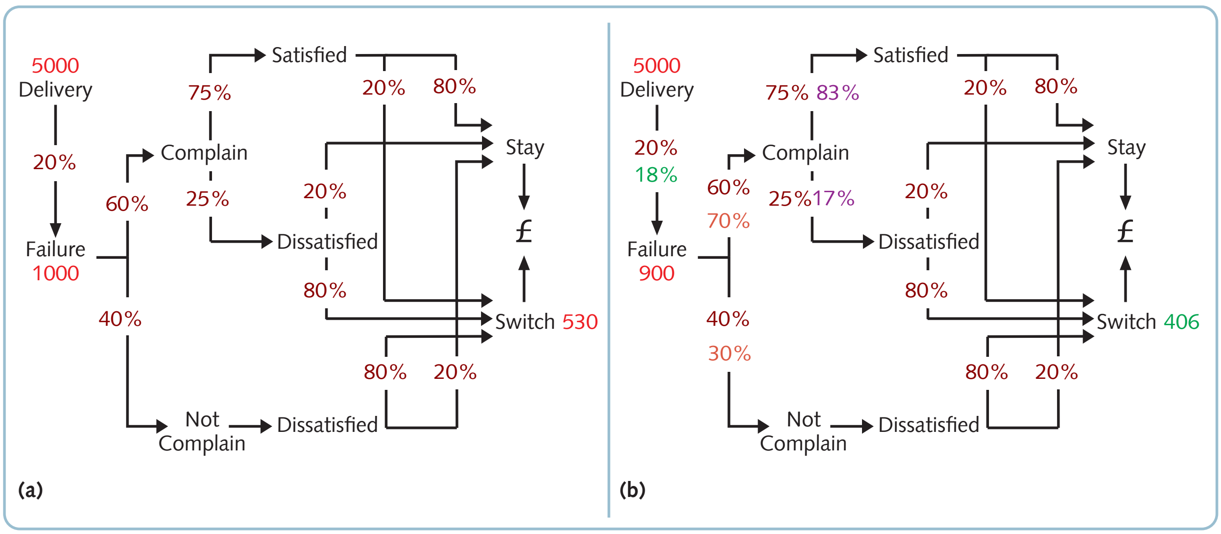


Figure 14.9 **Complaint value chain: (a) initial value chain and (b) with small improvements to each step**

example 80 per cent to 20 per cent) while satisfied complainers will tend to remain loyal (again, in this case 80 per cent to 20 per cent). So, assuming these percentages are correct, for every 5,000 customers processed by this particular service operation, 530 will switch.

Now let's assume that the operations manager decides to invest in small improvements to all stages in the complaint value chain. In Figure 14.9(b) the company has reduced its failures from 20 per cent to 18 per cent (still very poor of course) and has encouraged more customers who experienced a failure to come forward and complain. So the percentage complaining has risen from 60 per cent to 70 per cent. It has also made sure that a higher proportion (in this case, from 75 per cent to 83 per cent) of those who do make the effort to complain are satisfied. The end result is that the number of lost customers falls from 530 to 406. Assuming that an extra 124 customers have received value that is equal to, or more than, the costs of improvements, the organisation is making a good investment in its recovery and prevention efforts. What is important to understand here is how a relatively small improvement across the failure and complaint process can have such a significant impact on customer loyalty and switching.

The recovery process

Recovery needs to be a planned process. Organisations therefore need to design appropriate responses to failure, linked to the cost and the inconvenience caused by the failure to their customers. Such recovery processes need to be carried out either by empowered front-line staff or by trained personnel who are available to deal with recovery in a way that does not interfere with day-to-day service activities. Figure 14.10 illustrates a typical recovery sequence.

Discover

The first thing any manager needs to do when faced with a failure is to discover its exact nature. Three important pieces of information are needed: first of all, what exactly has happened; second, who will be affected by the failure; and third, why did the failure occur? This last point is not intended to be a detailed inquest into the causes of failure (that comes later) but it is often useful to know something of the causes of failure in case it is necessary to determine what action to take.

Act

The discover stage could take only minutes, or even seconds, depending on the severity of the failure. If the failure is a severe one with important consequences, we need to move on to doing

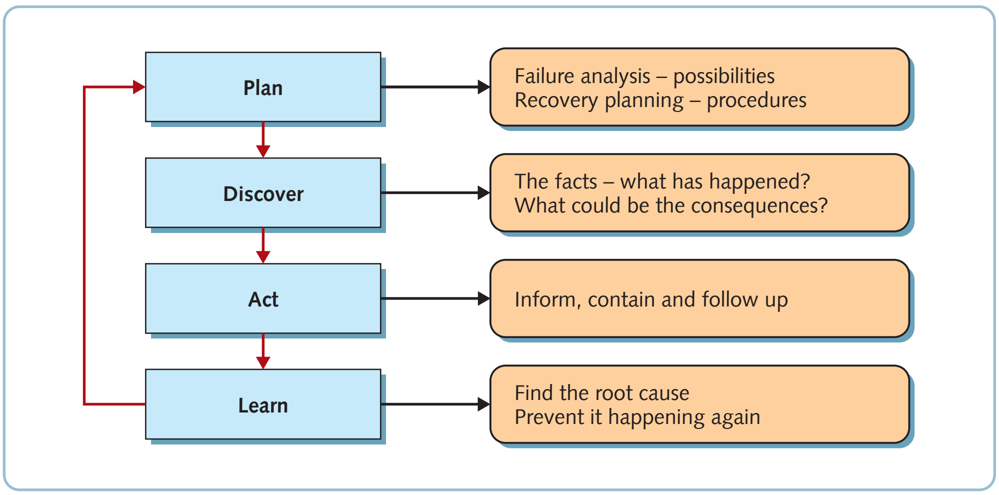


Figure 14.10 Recovery sequence for minimising the impact from failure

something about it quickly. This means carrying out three actions, the first two of which could be carried out in reverse order, depending on the urgency of the situation. First, tell the significant people involved what you are proposing to do about the failure. In service operations, this is especially important where the customers need to be kept informed. In all operations, however, it is important to communicate what action is going to happen so that everyone can set their own recovery plans in motion. Second, the effects of the failure need to be contained in order to stop the consequences spreading and causing further failures. The precise containment actions will depend on the nature of the failure. Third, there needs to be some kind of follow-up to make sure that the containment actions really have contained the failure.

Learn

The benefits of failure in providing learning opportunities should not be underestimated. In failure planning, learning involves revisiting the failure to find out its root cause and then engineering out the causes of the failure so that it will not happen again.

Plan

Learning the lessons from a failure is not the end of the procedure. Operations managers need formally to incorporate the lessons into their future reactions to failures. This is often done by working through 'in theory' how they would react to failures in the future.

Critical commentary

- Using the 'black swan' metaphor to think about low-probability but high-impact events can be seen as presenting a danger. That is, that we are tempted to treat such events as so rare that we fail to prepare for them. This issue was much debated during the COVID-19 pandemic. Some commentators believed that those accountable for the preparation for such an occurrence too easily dismissed their failures because of the perceived exceptional nature of the event. Also, the idea that failure can be detected through in-process inspection is seen increasingly as only partially true. Although inspecting for failures is an obvious first step in detecting them, it is not even close to being 100 per cent reliable. Accumulated evidence from research and practical examples consistently indicates that people, even when assisted by technology, are not good at detecting failure and errors. This applies even when special attention is being given to inspection. No one is advocating abandoning inspection as a failure detection mechanism. Rather it is seen as one of a range of methods of preventing failure. Much of the previous discussion surrounding the prevention of failure has assumed a 'rational' approach. In other words, it is assumed that operations managers and customers alike will put more effort into preventing failures that are either more likely to occur, or more serious in their consequences. Yet this assumption is based on a rational response to risk. In fact, being human, managers often respond to the perception of risk rather than its reality.

SUMMARY CHECKLIST

- ☐ Does the business have an operations and process resilience policy?
- ☐ Have all potential sources of failure been identified?
- ☐ Have any future changes in the sources of failure been identified?
- ☐ Is post-failure analysis carried out when failure does occur?
- ☐ Has the likelihood of each potential failure been assessed?
- ☐ Has the possibility of non-evident failures been addressed?
- ☐ Has the impact of all potential sources of failure been assessed?

- ☐ Are techniques such as failure mode and effect analysis (FMEA) used?
- ☐ Has due attention been paid to the possibility of designing out failure points?
- ☐ Has the idea of fail-safeing (*poka yoke*) been considered as a means of reducing the likelihood of failure?
- ☐ Have all approaches to process and technology maintenance been explored?
- ☐ Is the concept of redundancy economically viable for any potential failures?
- ☐ Does the operation have a failure mitigation plan?
- ☐ Has the whole range of mitigation actions been thoroughly evaluated?
- ☐ Are specific plans in place for the use of each type of mitigation action?
- ☐ Is a well-planned recovery procedure in place?
- ☐ Does the recovery procedure cover all the steps of discover, act, learn and plan?

Case study

Slagelse Industrial Services

Slagelse Industrial Services (SIS) had become one of Europe's most respected suppliers of die-cast zinc, aluminium and magnesium parts to hundreds of companies in many industries, especially automotive and defence. The company cast and engineered precision components by combining the most modern production technologies with precise tooling and craftsmanship. Slagelse Industrial Services (SIS) began life as a classic family firm under Erik Paulsen, who opened a small manufacturing and die-casting business in his home town of Slagelse, a town in east Denmark, about 100 km south-west of Copenhagen. He had successfully leveraged his skills and passion for craftsmanship over many years, while serving a variety of different industrial and agricultural customers. His son Anders had spent nearly ten years working as a production engineer for a large automotive parts supplier in the UK, but eventually returned to Slagelse to take over the family firm. Exploiting his experience in mass manufacturing, Anders spent years building the firm into a larger-scale industrial component manufacturer, but retained his father's commitment to quality and customer service. After 20 years, he sold the firm to a UK-owned industrial conglomerate and within ten years it had doubled in size again and now employed in the region of 600 people and had a turnover approaching £200 million. Throughout this period, the firm had continued to target its products into niche industrial markets where its emphasis upon product quality and dependability meant it was less vulnerable to price and cost pressures. However, in 2009, in the midst of difficult economic times and widespread industrial restructuring, it had been encouraged to bid for higher-volume, lower-margin work. This process was not very successful but eventually culminated in a tender for the design and production of a core metallic element of a child's toy (a 'transforming' robot).

Interestingly, the client firm, Alden Toys, was also a major customer for other businesses owned by SIS's corporate parent. It was adopting a preferred supplier policy and intended to have only one or two purchase points for specific elements in its global toy business. It had a high degree of trust in the parent organisation and on visiting the SIS site was impressed by the firm's depth of experience and commitment to quality. In 2010, it selected SIS to complete the design and begin trial production.

'*Some of us were really excited by the prospect . . . but you have to be a little worried when volumes are much greater than anything you've done before. I guess the risk seemed okay because in the basic process steps, in the type of product if you like, we were making something that felt very similar to what we'd been doing for many years*.' (SIS operations manager)

'*Well obviously we didn't know anything about the toy market but then again we didn't really know all that much about the auto industry or the defence sector or any of our traditional customers before we started serving them. Our key competitive advantage, our capabilities, call it what you will, they are all about keeping the customer happy, about meeting and sometimes exceeding specification*.' (SIS marketing director)

The designers had received an outline product specification from Alden Toys during the bid process and some further technical detail afterwards. Upon receipt of this final brief, a team of engineers and managers confirmed that the product could and would be manufactured using an up-scaled version of current production processes. The key operational challenge appeared to be accessing sufficient (but not too much) capacity. Fortunately, for a variety of reasons, the parent company was very supportive of the project and promised to underwrite any sensible capital expenditure plans. Although this opinion of the nature of the production challenge was widely accepted throughout the firm (and shared by Alden Toys and SIS's parent group) it was left to one specific senior engineer to actually sign both the final bid and technical completion documentation. By early 2011, the firm had begun a trial period of full-volume production. Unfortunately, as would become clear later, during this design validation process SIS had effectively sanctioned a production method that would prove to be entirely inappropriate for the toy market, but it was not until 12 months later that any indication of problems began to emerge.

Throughout both North America and Europe, individual customers began to claim that their children had been 'poisoned' while playing with the end product. The threat of litigation was quickly levelled at Alden Toys and the whole issue rapidly became a 'full-blown' child health scare. A range of pressure groups and legal damage specialists supported and acted to aggregate the individual claims. Although similar accusations had been made before, the litigants and their supporters focused in on the recent changes made to the production process at SIS and in particular the role of Alden Toys in managing its suppliers.

'*. . . it's all very well claiming that you trust your suppliers but you simply cannot have the same level of control over another firm in another country. I am afraid that this all comes down to simple economics, that Alden Toys put its profits before children's health. Talk about trust . . . parents trusted this firm to look out for them and their families and have every right to be angry that boardroom greed was more important!*' (Legal spokesperson for US litigants, being interviewed on a UK TV consumer rights show)

Under intense media pressure, Alden Toys rapidly convened a high-profile investigation into the source of the contamination. It quickly revealed that an 'unauthorised' chemical had been employed in an apparently trivial metal cleaning and preparation element of the SIS production process. Although when interviewed by the US media, the parent firm's legal director emphasised there was '*no causal link established or any admission of liability by either party*', Alden Toys immediately withdrew its order and began to signal an intent to bring legal action against SIS and its parent. This action brought an immediate end to production in this part of the operation, and the inspection (and subsequent official and legal visits) had a crippling impact upon the productivity of the whole site. The competitive impact of the failure was extremely significant. After over a year of production, the new product accounted for more than a third (39 per cent) of the factory's output. In addition to major cash-flow implications, the various investigations took up lots of managerial time and the reputation of the firm was seriously affected. As the site operations manager explained, even its traditional customers expressed concerns.

'*It's amazing, but people we had been supplying for thirty or forty years were calling me up and asking "[Manager's name] what's going on?" and that they were worried about what all this might mean for them . . . these are completely different markets!*'

Questions

1. What operational risks did SIS face when deciding to become a strategic supplier for Alden Toys?
2. What control problems did it encounter in implementing this strategy (pre- and post-investigation)?

Applying the principles

Some of these exercises can be answered by reading the chapter. Others will require some general knowledge of business activity and some might require an element of investigation. **Model answers for the first two questions are available on the student companion website**. *For model answers to the other questions in this section, please ask your tutor/lecturer.*

1. Revisit the examples of failure described in this chapter (Deepwater Horizon, the COVID-19 pandemic, 'Dieselgate', TSB's IT meltdown). Compile a table that summarises your view of:

a) The reasons for the main failure.

b) How it might have been prevented.

c) The result (damage) of the failure.

2. Wyco is a leading international retailer selling clothing and accessories, with stores throughout the United States, Europe and the Middle East. The countries from which it sources its products include Sri Lanka, Bangladesh, India and Vietnam. It was shocked when a British newspaper reported that an unauthorised subcontractor had used child workers to make some of its products at a factory in Delhi. In response, Wyco immediately issued a statement.

'*Earlier this week, the company was informed about an allegation of child labor at a facility in India that was working on one product for Wyco. An investigation was immediately launched. The company noted that a very small portion of a particular order placed with one of its vendors was apparently subcontracted to an unauthorised subcontractor without the company's knowledge or approval. This is in direct violation of the company's agreement with the vendor under its Code of Vendor Conduct*.' The company CEO said, '*We strictly prohibit the use of child labor. This is a non-negotiable for us – and we are deeply concerned and upset by this allegation. As we've demonstrated in the past, Wyco has a history of addressing challenges like this head-on. Wyco ceased business with 20 factories due to code violations. We have 90 people located around the world whose job is to ensure compliance with our Code of Vendor Conduct. As soon as we were alerted to this situation, we stopped the work order and prevented the product from being sold in stores. While violations of our strict prohibition on child labor in factories that produce product for the company are extremely rare, we have called an urgent meeting with our suppliers in the region to reinforce our policies. Wyco has one of the industry's most comprehensive programs in place to fight for workers' rights. We will continue to work with stakeholder organizations in an effort to end the use of child labor*.'

a) '*Being an ethical company isn't enough any more. These days, leading brands are judged by the company they keep*.' What does this statement mean for Wyco?

b) When Wyco found itself with this supply chain problem, did it respond in the right way?

3. One cause of aircraft accident is 'controlled flight into ground'. Predominantly, the reason for this is not mechanical failure but human failure, such as pilot fatigue. Boeing, which dominates the commercial airline business, has calculated that over 60 per cent of all the accidents that have occurred in the past 10 years had flight crew behaviour as their 'dominant cause'. For this type of failure to occur, a whole chain of minor failures must happen. First, the pilot at the controls has to be flying at the wrong altitude – there is only one chance in a thousand of this. Second, the co-pilot would have to fail to cross-check the altitude – only one chance in a hundred of this. The air traffic controllers would have to miss the fact that the plane was at the wrong altitude (which is not strictly part of their job) – a one-in-ten chance. Finally, the pilot would have to ignore the ground proximity warning alarm in the aircraft (which can be prone to giving false alarms) – a one-in-two chance.

a) What are your views on the quoted probabilities of each failure described above occurring?

b) How would you try to prevent these failures occurring?

c) If the probability of each failure occurring could be reduced by a half, what would be the effect on the likelihood of this type of crash occurring?

4. (True story) The light bulb in the men's lavatories of a factory finally burnt out after 70 years of operation. The manager at the firm said, '*It is actually a little bit sad. I joined the firm when I was fifteen (he is now sixty three) and it was there then*'. In fact, the bulb had survived bombs dropped in the Second World War that had devastated neighbouring buildings, the army firing its guns in the next-door park and punk band 'The Clash' playing at a neighbouring venue, which caused residents to complain that their windows were being shaken by the noise. When it did eventually fail, the firm had it mounted on a stand and gave it a place of honour. More remarkable, they tracked down the original supplier of the bulb. He had also survived. He was 99 years old and still remembered selling them. Does this incident invalidate the use of failure data in estimating component life?

5. An automated sandwich-making machine in a food manufacturer's factory has six major components, with individual reliabilities as shown in Table 14.2.

   a) What is the reliability of the whole system?

   b) If it is decided that the wrapper in the automated sandwich-making machine is too unreliable and a second wrapper is needed that will come into action if the first one fails, what will happen to the reliability of the machine?

Table 14.2 Individual reliabilities of major components

Component	*Reliability*
Bread slicer	0.97
Butter applicator	0.96
Salad filler	0.94
Meat filler	0.92
Top slice of bread applicator	0.96
Wrapper	0.91

Notes on chapter

1 According to Nassim Nicholas Taleb, the theorist who coined the term, the three characteristics of a black swan event are rarity, extreme impact and retrospective predictability. It is based on an old saying that black swans did not exist, until one appeared to prove otherwise.

2 Aldrick, P. (2020) 'Overcapacity is not a bad thing when it protects us from risk', *The Times*, 4 April.

3 Economist (2020) 'Supple supplies: Businesses are proving quite resilient to the pandemic', *Economist* print edition, 16 May.

4 Bourgeois, L.J. (1981) 'On the measurement of organizational slack', *The Academy of Management Review*, 6 (1), pp. 29–39.

5 The information on which this example is based is taken from: Borunda, A. (2020) 'We still don't know the full impacts of the BP oil spill, 10 years later', *National Geographic*, 20 April; Hill, A. (2020) 'Covid-19 lays bare managers' efficiency obsession', *Financial Times*, 20 April; Economist (2020) 'Supple supplies: Businesses are proving quite resilient to the pandemic', *Economist* print edition, 16 May.

6 See https://www.telegraph.co.uk/travel/news/Heathrow-seeks-high-wire-walkers-to-change-light-bulbs/ [accessed 8 October 2020].

7 The information on which this example is based is taken from: Amelang, S. and Wehrmann, B. (2020) '"Dieselgate" – a timeline of the car emissions fraud scandal in Germany', Factsheet, Clean Energy Wire, 25 May, https://www.cleanenergywire.org/factsheets/dieselgate-timeline-car-emissions-fraud-scandal-germany [accessed 2 October 2020]; Tovey, A. (2017) 'VW attacked by MPs over failure to release findings of "dieselgate" investigation', *Telegraph*, 22 March.

8 Takala, J., Hämäläinen, P., Saarela, K.L. et al. (2014) 'Global estimates of the burden of injury and illness at work in 2012', *Journal of Occupational and Environmental Hygiene*, 11 (5), pp. 326–337.

9 The information on which this example is based is taken from: Makortoff, K. (2019) 'TSB lacked common sense in run-up to IT meltdown, says report', *Guardian*, 19 November; Megaw, N. (2018) 'TSB swings into red in first half following IT debacle', *Financial Times*, 27 July.

10 Justin Fier, quoted in Walker, M. (2020) 'Darktrace: An AI cybersecurity platform that serves as the immune system for enterprise business data by fighting off threats', Cardrates, 3 February, https://www.cardrates.com/news/darktrace-is-an-ai-based-enterprise-immune-system/ [accessed 2 October 2020].

11 The information on which this example is based is taken from: Darktrace website, https://www.darktrace.com/en/ [accessed 2 October 2020]; Walker, M. (2020) 'Darktrace: An AI cybersecurity platform that serves as the immune system for enterprise business data by fighting off threats', Cardrates, 3 February, https://www.cardrates.com/news/darktrace-is-an-ai-based-enterprise-immune-system/ [accessed 2 October 2020]; Ismail, N. (2019) 'Darktrace unveils the cyber AI analyst: A faster response to threats', Information Age, 4 September; Ross, A. (2019) 'ML and AI in cyber security: Real opportunities overshadowed by hype', Information Age, 7 March.

12 The information on which this example is based is taken from: Economist (2019) 'An emergency landing system that passengers can activate', *Economist* print edition, 28 November.

Taking it further

Hopkin, P. ***(2017) Fundamentals of Risk Management: Understanding, evaluating and implementing effective risk management, 4th edition, Kogan Page.*** *A comprehensive introduction to risk with good coverage of many core frameworks.*

Hubbard, D.W. ***(2009) The Failure of Risk Management: Why it's broken and how to fix it, John Wiley & Sons.*** *An interesting read, particularly for those who like the critical commentaries in this text! A polemic, but one that is clearly written.*

Smith, D.J. ***(2011) Reliability, Maintainability and Risk, 8th edition, Butterworth-Heinemann.*** *A comprehensive and excellent guide to all aspects of maintenance and reliability. The book has a good mix of qualitative and quantitative perspective on the subject.*

Waters, D. ***(2011) Supply Chain Risk Management: Vulnerability and resilience in logistics, 2nd edition, Chartered Institute of Logistics and Transportation/Kogan Page.*** *Provides a very detailed and practical guide to considering risks within operations and supply chains.*

15 Project management

Introduction

In this chapter, we examine how projects of all shapes and sizes can be executed more successfully. To do this, managers must first understand the innate characteristics of the project, as well as the implications of differences between projects. Second, they must appreciate the vital role of effective project management in influencing the success (or failure!) of projects and recognise the key responsibilities and skills of those tasked with running projects. Third, they must understand the environment within which their project is being undertaken and determine how best to manage project stakeholders. Fourth, they must effectively define and plan projects while balancing competing performance objectives of quality, time and cost. Finally, they must effectively control projects through their life cycle and ensure that learning between projects is maximised. Figure 15.1 shows where project management fits within the overall structure of this text.

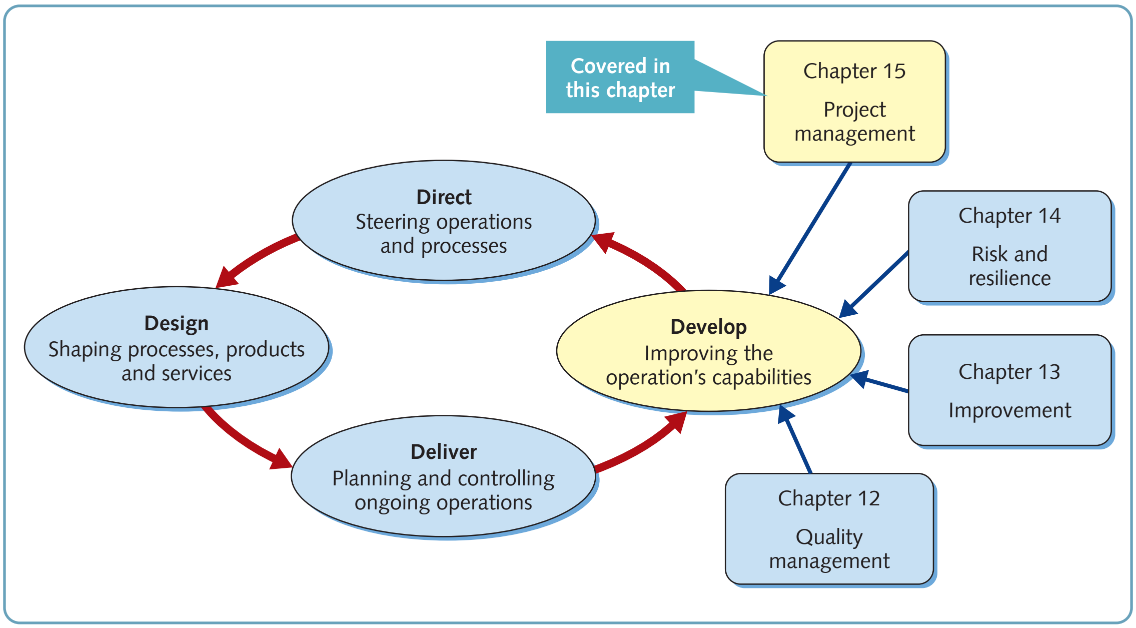


Figure 15.1 **Managing projects is the activity of defining, planning, controlling and learning from projects**

EXECUTIVE SUMMARY

15.1 Are the innate characteristics of the project understood?
↓
15.2 Is the applicability of project management understood?
↓
15.3 Are the project environment and project stakeholders understood?
↓
15.4 Has the project been effectively defined?
↓
15.5 Has the project been adequately planned?
↓
15.6 Is the project effectively controlled and learned from?

15.1 Are the innate characteristics of the project understood?

A project is a temporary activity aimed at achieving a specific and highly customised goal, within a set time frame, using a defined group of resources. They involve many non-routine and complex tasks, which means projects are often highly uncertain. There are many different types of project, including infrastructure or capital investment, new product or service development, events, and projects focused on delivering organisational change. The process of managing a project involves not only understanding their common characteristics, but also understanding key differences from one project to the next. In this chapter, we focus on differences in the level of (a) innovation, (b) time pressure and (c) complexity of the project being managed.

15.2 Is the applicability of project management understood?

Project management is the activity of defining, planning, controlling and learning from projects. Beyond this 'life cycle' perspective, it is also concerned with effectively balancing deliverables (quality), time and cost objectives within the so-called 'iron triangle'. Finally, from an organisational perspective, project management involves coordinating these life cycles and competing objectives across multiple functions. Project managers are tasked with organising the project team, assigning responsibilities, working with stakeholders and running the project on a daily basis. This position is a highly skilled one, requiring technical project management knowledge, interpersonal skills and leadership ability. Very often, project managers need the ability to motivate staff who not only report into managers other than themselves, but also divide their time between several different projects. It is also important to note that project management is not just the province of project managers. It is a ubiquitous task of operations management, because people in all types of operations will get involved in some element of project management.

15.3 Are the project environment and project stakeholders understood?

Projects do not exist in a vacuum. It is vital that the project team understand the key characteristics of the environment within which their project is being undertaken.

The project environment includes the internal environment, business environment, economic–political environment and geo-social environment. The project team must determine how best to engage those individuals, groups or entities that affect, or are affected by, the project – a process called stakeholder management. This activity is based on three basic activities – identifying stakeholders, understanding their different perspectives and managing different, and often competing, interests.

15.4 Has the project been effectively defined?

Defining a project involves three related activities – setting project objectives, scoping the project and developing a project strategy. Most projects can be defined by the relative importance of three objectives. These are cost – keeping the overall project to its original budget, time – finishing the project by the scheduled finish time, and quality – ensuring that the project outcome is as was originally specified. The project scope identifies its work content and its outcomes. The task is concerned not only with what the project will do, but also what it will not do – its limits and exclusions. The project strategy describes the general way in which the project is going to meet its objectives, including significant project milestones and 'stagegates'.

15.5 Has the project been adequately planned?

Project planning involves more detailed analysis to help determine the cost and duration of the project and the level of resources that will be required. In this chapter, we outline five stages of project planning – (1) identifying activities, (2) estimating time and resources, (3) identifying relationships and dependencies between activities, (4) identifying time and resource schedule constraints, and (5) fixing the final schedule. Critical path analysis (CPA) and programme evaluation and review technique (PERT) are often used to aid the project planning process.

15.6 Is the project effectively controlled and learned from?

Project control and learning involves monitoring the project in order to check its progress, assessing performance against project plans, intervening in projects, and using learning to improve subsequent projects. Common forms of monitoring include assessing deviations in planned spending, the number and length of delays, and flagging changes to technical deliverables. In assessing project performance, earned value analysis (EVA) is a widely adopted method that allows for comparisons between expected costs and schedules and the actual performance of a project. Project intervention typically occurs when a project is out of control in relation to one or more aspects of time, cost and quality. It may also involve deciding when to devote extra resources to accelerate individual activities within the project – an activity typically known as 'crashing'. Project learning is essential in developing longer-term organisational capability to manage projects more effectively as we move from one project to the next. This requires a shift towards higher levels of formalised learning being embedded within the end-to-end project process.

15.1 Diagnostic question: *Are the innate characteristics of the project understood?*

What are the characteristics of projects?

A project is a set of activities that must be completed to deliver a specific goal within a set time frame, using a defined group of resources. The process of managing a project begins by understanding its fundamental characteristics. Projects share a number of common features in being:

- temporary – while some projects last hours and others many years, they all have a defined start and end point;
- dedicated to completing a specific goal within key time, cost and quality requirements;
- something with an outcome that is typically unique or at least highly customised;
- based on many non-routine and complex tasks.

It is these features that generate high levels of risk and uncertainty in the management of projects, and why changed specifications (quality), severe delays (time), cost escalation (cost) and major disputes between key stakeholders are commonplace.

Projects come in all shapes and sizes. Some are associated with the development of specific services or products. Some are capital investment or infrastructure projects, such as organising the installation of new buildings or process technologies. Some involve specific pre-planned organisational changes that seek to make a change to methods of working or reporting responsibilities. Some are large and complex with multiple stakeholders, such as a major public infrastructure initiative, while others are short, small-scale and limited, such as a simple event. Technically, many small-scale activities, taking minutes or hours, conform to the common features of projects described above. However, in this chapter we will focus mainly on relatively larger-scale projects, lasting months, years or even several decades.

It is also worth pointing out the distinction between 'projects' and 'programmes'. A programme, such as a continuous improvement programme, has no defined end point. Rather it is an ongoing process of change. Individual projects, such as the development of a new training workshop, may be individual sub-sections of the overall programme. Programme management will overlay and integrate the individual projects. Generally, this is a more difficult task in the sense that it requires resource coordination, particularly when multiple projects share common resources.

Case example

'For the benefit of all' – NASA's highs and lows[1]

Space remains one of the most challenging contexts for carrying out successful projects, and since humankind first ventured into space we have witnessed many project failures. In January 1986, NASA's (National Aeronautics and Space Administration) Space Shuttle *Challenger* exploded shortly after take-off, killing all on board. The failure was traced to a faulty seal, which ruptured and caused the liquid hydrogen fuel to explode. The NOAA weather satellite was NASA's first mission after a suspension of 32 months to allow investigation of the disaster. Things did not go well – just 71 seconds into its flight, the launch rocket was struck by lightning. With its first-stage rockets disabled, ground control destroyed the rocket to minimise the risk of it falling back to Earth.

Project failures, however, have often enabled future successes for NASA. Take the Apollo 6 mission, the final non-manned test for the Saturn V rocket, in April 1968. Shortly after launching, the rocket experienced 'pogo oscillations' (variations in thrust levels caused by changing fuel rates), two engines shut down prematurely in the second-stage burn, and the third-stage rocket failed to reignite. Yet, in July 1969, the same (improved!) rocket successfully delivered the *Apollo* Lunar Module into space on the manned *Apollo 11* mission that delivered Neil Armstrong and Edwin ('Buzz') Aldrin safely onto the surface of the Moon.

On other occasions problems have occurred, but a combination of the technical capabilities and inventiveness of the key stakeholders have prevented project failure. For example, the *Apollo 13* mission in 1970 was intended to be the third to land on the Moon. However, following a failure in the service module's oxygen tank two days into the mission, the new project objective became the safe return of the crew back to Earth. This involved a move from the service module to the lunar module as a form of 'lifeboat', followed by various improvisations to convert a craft that was originally designed to support two men on the lunar surface for two days, to one able to support three men in space for four days. Fortunately, the revised plan worked and while the mission failed in its primary objectives, it was very successful in its revised scope.

In another example, the Hubble Space Telescope had many issues post launch, including problems focusing the telescope due to errors made by European and US scientists in translating the units of measurement! However, the project team found ways to overcome the various failures and the telescope now provides some of the most detailed images of deep space – as well as a significantly expanded understanding of our (tiny) place in the universe.

Some space projects have gone far better than expected, leading to change in scope due to their success. For example, the Cassini mission (a collaboration involving NASA, the European Space Agency and the Italian Space Agency) began with the objective to reach Saturn, some 750 million miles away. On its way, Cassini took photographs of our solar system, including fly-bys of Earth, Venus and Jupiter – the photos of Jupiter being the most detailed ever taken of the planet. Cassini also confirmed Einstein's theory of general relativity while travelling to Saturn. On arrival, Cassini successfully deployed the Huygens probe, which began returning data to Earth from Saturn's largest moon, Titan. Meanwhile, Cassini continued to collect detailed data and images on the planet and its other moons. The mission, originally expected to last four years, was first extended by two years (the Cassini Equinox Mission) and then again for a further seven years (the Cassini Solstice Mission) as the spacecraft continued to function effectively. It wasn't until 2017, nine years after the planned project end, that Cassini was finally 'de-orbited' to burn up in Saturn's atmosphere, though not before it had completed a number of high-risk passes within Saturn's inner rings to maximise its total scientific contribution.

It's not just NASA and its collaborators that have looked to build on previous experiences to improve their projects. The recent generation of commercial space travel and exploration companies, such as SpaceX, Sierra Nevada Corporation, Boeing, Northrop Grumman Innovation Systems, Blue Origin and Virgin Galactic, have all sought to gain insights from established space agencies. This not only offers the benefit of (hopefully) avoiding mistakes that have been made in the past, but has also enabled significant reductions in the costs of development, testing and operations. For example, SpaceX has made major breakthroughs in fuel, engines and, most valuably, in increasing the proportion of its rocket and launch vehicles that it can recover and reuse. The result is that the cost of launching a kilogram of material into space has fallen to (just!) $2,720 for the SpaceX *Falcon 9*, the rocket used to successfully deliver NASA astronauts to the International Space Station in 2020. This compares to an eye-watering $54,500 per kilogram when the NASA space shuttle programme was operational between 1981 and 2011.

JPL-Caltech/NASA

Differentiating between projects

OPERATIONS PRINCIPLE
The challenge of managing a project is contingent on its level of innovation, time pressure and complexity.

So far, we have shown what projects have in *common* – temporary activities, with specific and highly customised goals, involving many non-routine and complex tasks. However, it is also critical to understand *differences* between projects. Here, we focus on differences in the level of (a) innovation, (b) time pressure and (c) complexity of the project being managed. Figure 15.2 illustrates the profiles of four different projects using this form of differentiation.

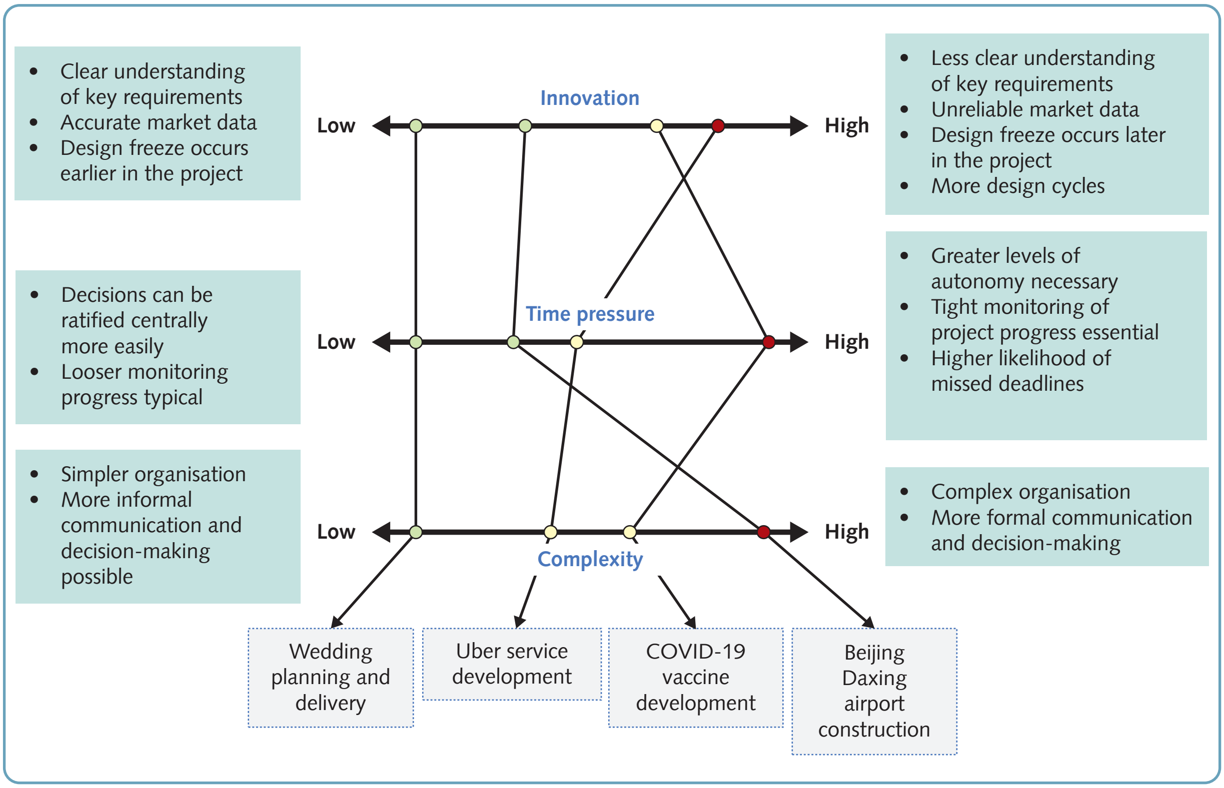


Figure 15.2 Differentiating projects based on the level of innovation, time pressure and complexity, and the implications of these differences

Level of project innovation

The first way to differentiate projects is to consider their relative level of innovation. For projects, innovation may entail the delivery of new services or products, the incorporation of new technologies, the development of new routes to market and the transformation of organisational processes, for example. Incremental projects typically involve relatively modest levels of innovation, building upon existing knowledge and/or resources where existing routines are not fundamentally changed. By contrast, radical projects exhibit high levels of novelty that require completely new knowledge and/or resources, and often make existing routines obsolete. Examples of these more innovative projects include the development and launch of Uber (the multinational ride-hailing service) and Airbnb, the rapid emergence of sensor technology in many aspects of farming, the first iPhone, which paved the way for the modern smartphone market, and the development of Netflix, with its huge impact on the home entertainment sector.

Level of project time pressure

The second way to differentiate projects is to consider the relative level of time pressure that they face. It's important to remember that time pressure is not about speed – some projects have urgency but last for many years, others are not urgent but last only a few weeks. Some projects face low levels of time pressure where the specific time frame is not deemed critical by the project stakeholders. Many public works and internal projects fall into this category. Some projects face moderate levels of time pressure where completion on time is important for competitive advantage and leadership. Many business-related projects, such as new service or product development, fall into this category. Finally, some projects face high levels of time pressure where there is a specific window of opportunity and any delay can mean project failure. For example, in May 1961, President John F. Kennedy delivered a speech to the US Congress in which he stated, '*I believe that this nation should commit itself to achieving the goal, before this decade is out, of landing a man on the moon and returning him safely to the earth*'. In doing so, he set a time frame that was to be critical to the ambitions of the moon-landing project. Other examples of projects facing high levels of time pressure are those created in response to specific crises such as the Ebola or COVID-19 virus pandemics, wars or natural disasters.

Level of project complexity

The third way to differentiate projects is to consider their relative level of complexity. Some projects exhibit low levels of complexity, are often self-contained and with a relatively small number of key stakeholders. Examples may include planning a wedding, creating an online sales platform, developing a new MBA operations and process management module or writing a new book(!). Other projects face greater levels of complexity, often combining a set of sub-elements and involving many more stakeholders. Examples include constructing a new research and development facility, developing a new portfolio of postgraduate education within a university or organising a large-scale music festival. While the sub-elements of the project have a common goal, the added complexity creates significantly higher coordination and integration challenges. Finally, some projects have to deal with extremely high levels of complexity, coordinating several major projects to deliver against a common goal. A good example of this kind of project is China's South-to-North water diversion project – a multi-decade infrastructural mega-project expected to be completed in 2050 at a cost of $71 billion.

15.2 Diagnostic question: *Is the applicability of project management understood?*

What is project management?

Project management is the activity of defining, planning, controlling and learning from projects. While this 'life cycle' perspective is useful and allows us to consider projects in a sequential manner, it is important to understand that project management is essentially an *iterative* process. Problems or changes that become evident in project control, for example, may require replanning and can even lead to changes in the original project scope. Going beyond the 'life cycle perspective', project management is also concerned with effectively balancing deliverables (quality), time and cost objectives within the so-called 'iron triangle'. Finally, from an organisational perspective, project management involves managing these life cycles and performance objectives across multiple functions within an organisation.

The activity of project management is very broad in that it could encompass almost all the operations management tasks described in this text. As such, it could have been treated almost anywhere within our direct, design, delivery, develop structure. We have chosen to place it in the context of operations development because the majority of projects that managers will be engaged in are essentially improvement projects. Of course, many projects are vast enterprises with very high levels of resourcing, complexity and uncertainty that will extend over many years. Such projects require professional project management involving high-level technical expertise and management skills. However, smaller projects that implement the many important improvements that will determine the strategic impact of operations development also benefit greatly from robust project management.

OPERATIONS PRINCIPLE
Project management is the activity of defining, planning, controlling and learning from projects.

Case example

McCormick's AI spice project[2]

Artificial intelligence (AI) is increasingly playing a transformative role in many aspects of business operations. Now we are witnessing its use in new product development projects within the food sector. McCormick, the largest spice company in the world, recently teamed up with IBM Research to develop an AI system aimed at developing new flavour combinations. The collaboration leverages IBM's expertise in machine learning and its proprietary IBM Research AI for Product Composition to sift through data on thousands of ingredients, sales (both its own and within the sector) and consumer taste trends, consumer testing information, and hundreds of thousands of existing seasoning mixes, to suggest potential new formulas. The system can also advise on possible substitutes for raw ingredients, relative level of novelty (based on the 'distance' between a flavour combination and its nearest neighbour) and likely human response.

The company has already launched its first products leveraging its new AI system – 'McCormick One' is a range of seasoning for simple one-dish recipes, including Tuscan Chicken, New Orleans Sausage and Bourbon Pork Tenderloin. The new AI system is in sharp contrast to McCormick's traditional approach to new product and service development (NPSD) projects, involving a large team of chefs, nutritionists, food scientists, chemists and chemical engineers building on 'seed formulas' to develop new flavour combinations. The company believe that the AI system can help in its attempts to develop more innovative spice mixes, partly by avoiding cultural biases that may be inherent within its human development team. For example, the system recently suggested that adding cumin to a pizza seasoning would improve the taste. Such a move had never been considered by McCormick's food scientists, but their subsequent consumer testing backed up the idea.

Not only does McCormick believe that its AI system can help generate more novel flavours, with a lower probability of market rejection, it also saves costs in product development and shortens the project time frames by up to two thirds. Some of the time is saved in the rapid creation of many different possible formulas, followed by automated filtering to create a shortlist of potential products for further human evaluation. Further time savings occur in the consumer testing phase of the projects, with feedback fed directly into the system, analysed, and then integrated into flavour revisions. In the highly competitive world of spice and flavourings, such time and cost savings offer substantial commercial benefits.

monticello/Shutterstock

Project managers and their skill sets

In order to coordinate the efforts of many people in different parts of the organisation (and often outside it as well), all projects need a project manager. The project manager is the person accountable for project delivery and she or he has a number of key responsibilities (see Figure 15.3). The project manager organises the project team, with the responsibility, if not always the authority, to run the project on a day-to-day basis. Many of a project manager's activities are concerned with managing human resources. The people working in the project team need a clear understanding of their roles in the (usually temporary) organisation. Controlling an uncertain project environment requires the rapid exchange of relevant information with the project stakeholders, both within and outside the organisation. People, equipment and other resources must be identified and allocated to the various tasks.

Undertaking these tasks successfully makes project management a very demanding role, requiring a diverse set of skills that include technical project management knowledge, interpersonal skills and leadership ability. Very often, project managers must motivate staff who not only report into managers other than themselves, but also divide their time between several

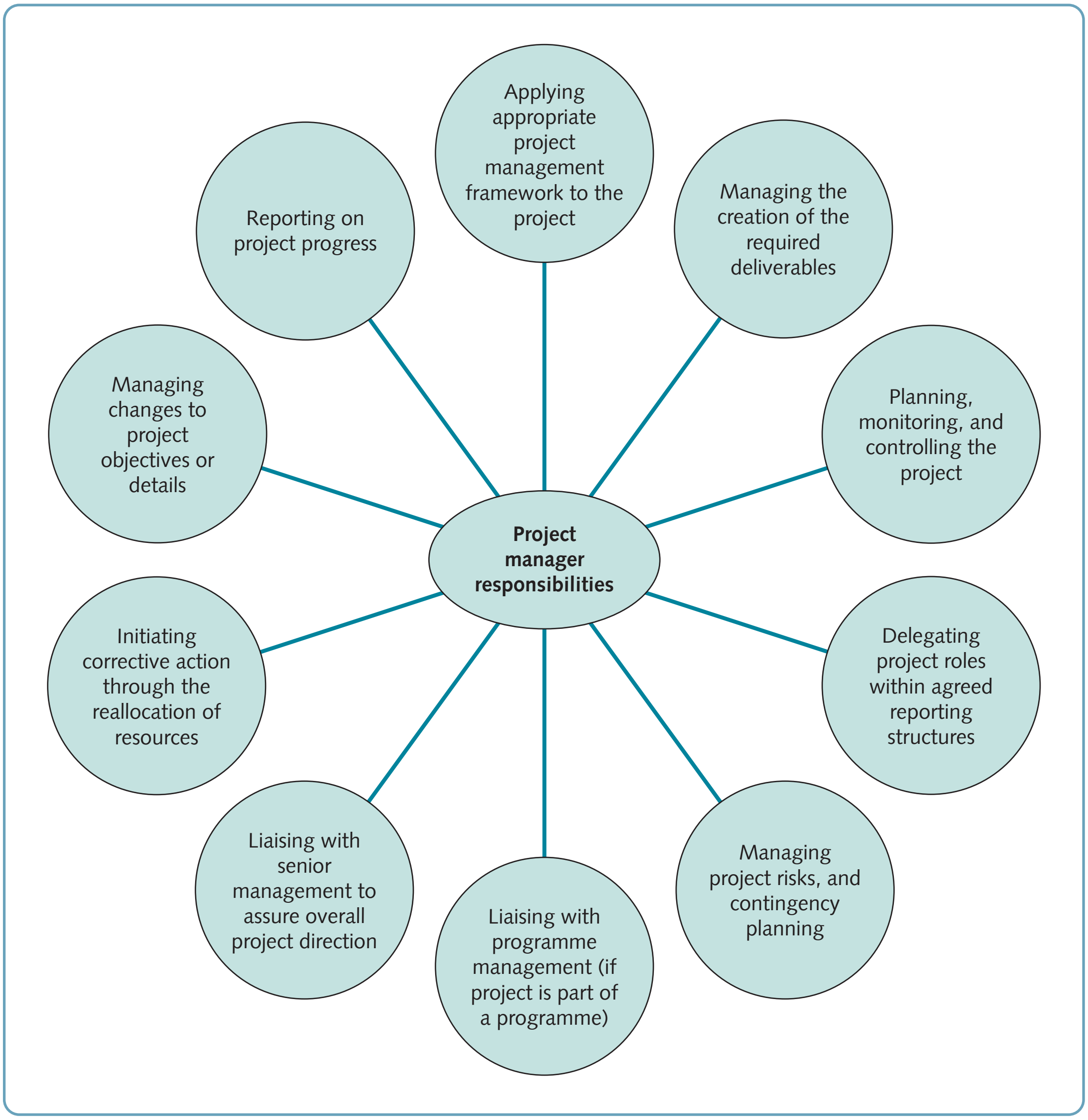


Figure 15.3 **Typical responsibilities of a project manager**

> OPERATIONS PRINCIPLE
> ***The activity of project management requires interpersonal as well as technical and leadership skills.***

different projects. In addition, they must pay attention to details without losing sight of the big picture, establish an open, communicative environment while remaining wedded to project objectives, and have an ability to remain optimistic while planning for the worst. Such challenges have led to the increasing professionalisation of project management over the last 20 years, with many more of those leading projects now holding professional qualifications.

Managing matrix tensions

In all but the simplest projects, project managers usually need to reconcile the interests of both the project itself and the departments contributing resources to the project. When calling on a variety of resources from various departments, projects are operating in a 'matrix management' environment, where projects cut across organisational boundaries and involve staff that are required to report to their own line manager as well as to the project manager. Figure 15.4 illustrates the type of reporting relationship that usually occurs in matrix management structures running multiple projects. A person in department 1 assigned part-time to projects A and B

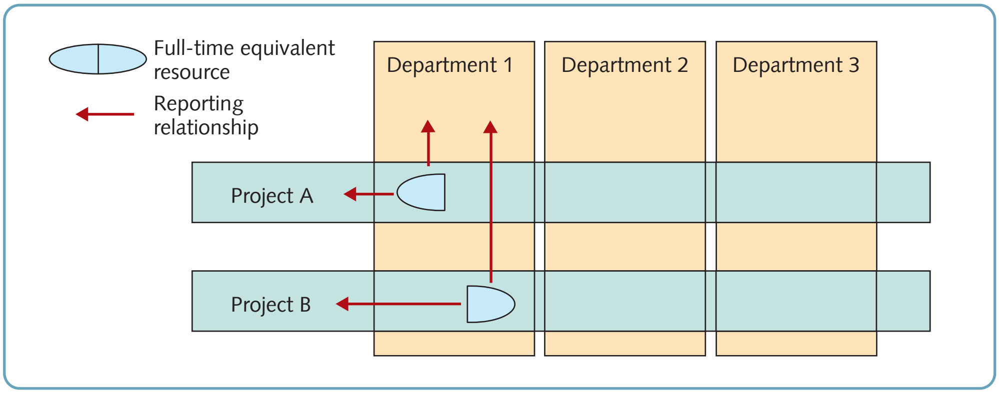


Figure 15.4 Matrix management structures often result in staff reporting to more than one project manager as well as their own department

will be reporting to three different managers, all of whom will have some degree of authority over their activities. This is why matrix management requires a high degree of cooperation and communication between all individuals and departments. Although decision-making authority will formally rest with either the project or departmental manager; most major decisions will need some degree of consensus. To function effectively, matrix management structures should have the following characteristics:

- Effective channels of communication between all managers involved, with relevant departmental managers contributing to project planning and resourcing decisions.
- Formal procedures in place for resolving any management conflicts that arise.
- Encouragement for project staff to feel committed to their projects as well as to their own department.
- Sufficient time devoted to planning the project and securing the agreement of the line managers to deliver on time and within budget.

Forming a project team and assigning responsibilities

A key role of the project manager is to form a project team and assign responsibilities for key tasks. In forming a team, it is important to consider the diversity of members to ensure that strengths and potential shortcomings of team members are balanced. For example, project teams need individuals who are naturally organised, with the capability to take ideas and make them work in practice. However, too many of this type of team member can limit flexibility and creativity within a project – a major limitation, particularly when working on projects requiring high levels of innovation. Conversely, projects with too many individuals who are predominantly free-thinking and creative often run into problems because such individuals are often less interested in the 'devil in the detail' of the project.

Once the project team is formed, the project manager needs to assign responsibility for all activities in the project (both to those within the project team and to third parties). A structured way to do this is in the form of a responsibility matrix. In its simplest form, this will simply identify who is responsible for each key activity in the project timeline. In some cases, responsibility matrices will show not only the responsible party but also others who are expected to provide support for each activity. The RACI matrix has emerged as a popular method of visualising responsibility, identifying those who are *responsible*, *accountable*, to be *consulted* and to be *informed*. Table 15.1 provides an example of a RACI matrix for a consulting project focused on new market testing in Chennai, India.

Table 15.1 RACI matrix for consulting project focused on new market testing in Chennai, India

Deliverable or task	*Ekta (Project Sponsor)*	*Jayesh (Project Manager)*	*Ritika (Technical Lead)*	*Punya (Analyst)*	*Shorya (Analyst)*	*Ashwin (Analyst*	*Shivani (Client)*
Phase 1 (scoping)							
Client kick-off meeting	A	R					C
Needs analysis	I	A	R				C
Analyst contracting	I	A	R				
Client review meeting	A	R	C				C
Phase 2 (data collection)							
Market research (focus group)		I	A	R			C
Market research (survey)		I	A		R		C
Market research (secondary data)		I	A			R	C
Phase 3 (analysis and report)							
Data analysis (focus group)	I	A	C	R			
Data analysis (survey)	I	A	C		R		
Data analysis (secondary data)	I	A	C			R	
Report (first draft)	I	A	R	C	C	C	
Client presentation	A	R	C				I
Report (final)	A	R	C				C
Project closure	R	C	I				A

[NOTE] R = Responsible; A = Accountable; C = Consult; I = Inform

15.3 Diagnostic question: *Are the project environment and project stakeholders understood?*

Projects do not exist in a vacuum. As such, it is vital that the project team understand the key characteristics of the environment within which their project is being undertaken and they identify the individuals, groups or entities that have an interest in the project process or outcome. They must then decide how to engage with these different stakeholders and how best to manage their competing needs.

Understanding the project environment

The project environment comprises all the factors that may affect the project during its life. Understanding the project environment is important because it affects the way in which a project will need to be managed and highlights the possible dangers that may cause the project to fail. Figure 15.5 illustrates four key aspects of the project environment.

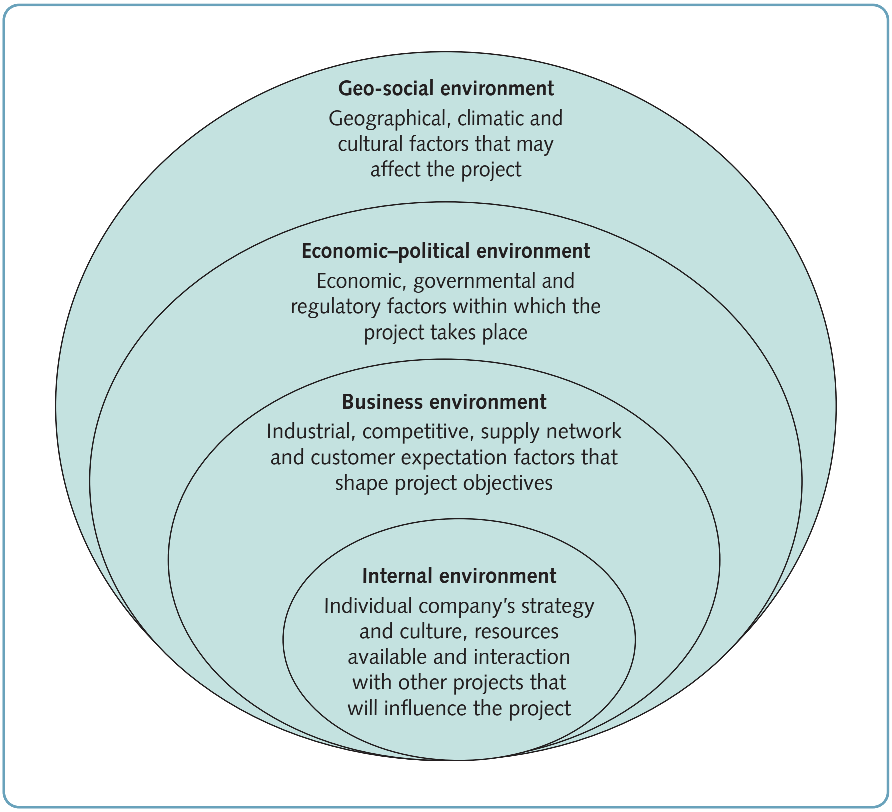


Figure 15.5 Understanding the elements of the project environment

Understanding and managing project stakeholders

Project stakeholders are those individuals, groups or entities that affect, or are affected by, the project. Internal stakeholders include the client, the project sponsor, the project team, functional managers, contractors and project support. External stakeholders (i.e. those outside of the core project, rather than necessarily outside of the organisation) include end users, suppliers, competitors, lobby groups, shareholders, government agencies and employees.

All projects will have stakeholders – complex projects will have many. They are likely to have different views on a project's objectives that may conflict with other stakeholders. As such, there is a practical incentive to include as many stakeholders as possible from an early stage, to prevent objections and problems later in the project. Moreover, there can be significant direct benefits from using a stakeholder-based approach. Project managers can use the opinions of powerful stakeholders to shape the project at an early stage. This makes it more likely that they will support the project and can also improve its quality. Communicating with stakeholders early and frequently can ensure that they understand the project and its potential benefits more fully. Stakeholder support may even help to win more resources, making it more likely that projects will be successful. Perhaps most importantly, managers can anticipate stakeholder reactions to various aspects of the project and plan the actions that could prevent (or deal with) opposition, or build support.

> OPERATIONS PRINCIPLE
> ***All projects involve stakeholders who may have different interests and priorities.***

Managing stakeholders can be a subtle and delicate task, requiring significant social and, sometimes, political skills. It is based on three basic activities – identifying stakeholders, understanding their different perspectives and managing them.

Identifying project stakeholders

Think of all the individuals, groups or entities who affect or are affected by your work, who have influence or power over it, or have an interest in its successful or unsuccessful conclusion. Figure 15.6 illustrates a stakeholder map for a technology platform project in the third sector (also known as the not-for-profit sector or charity sector) aimed at matching charities with funding opportunities. Although stakeholders are not just individuals, ultimately you must communicate with people; project managers should look to identify key individuals within a stakeholder organisation. In addition, even if one decides not to attempt to manage every identified stakeholder, the process of stakeholder mapping is still useful because it gets those working on a project to see the variety of competing forces that are often at play.

Understanding project stakeholders

Once all stakeholders have been identified, it is important to understand their different perspectives on the project. Some key questions that can help understand project stakeholders include:

- What financial or emotional interest do they have in the outcome of the project? Is it positive or negative?
- What motivates them most of all?
- What information do they need?
- What is the best way of communicating and consulting with them?
- What is their current opinion of the project?
- Who influences their opinions? Do some of these influencers therefore become important stakeholders in their own right?
- If they are not likely to be positive about the project, what will win them around?
- If you don't think they will support you, how might you manage their opposition?

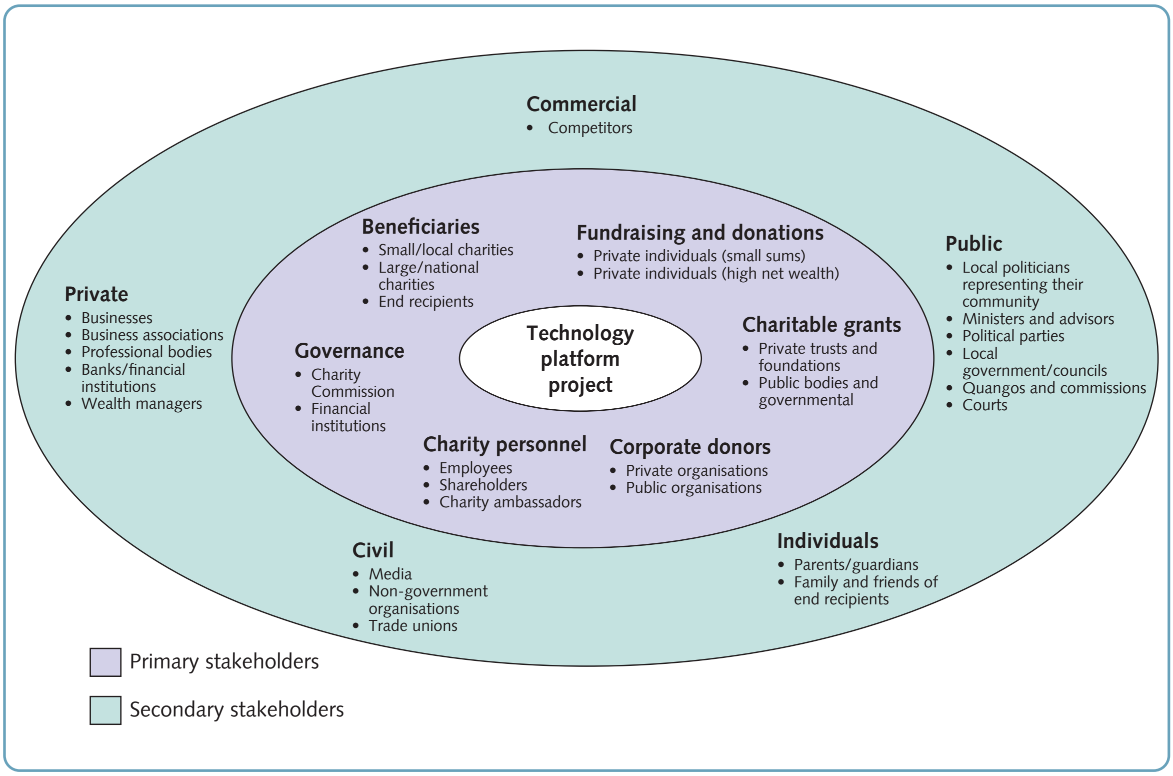


Figure 15.6 **Stakeholder mapping for a third-sector (not-for-profit) technology platform project**

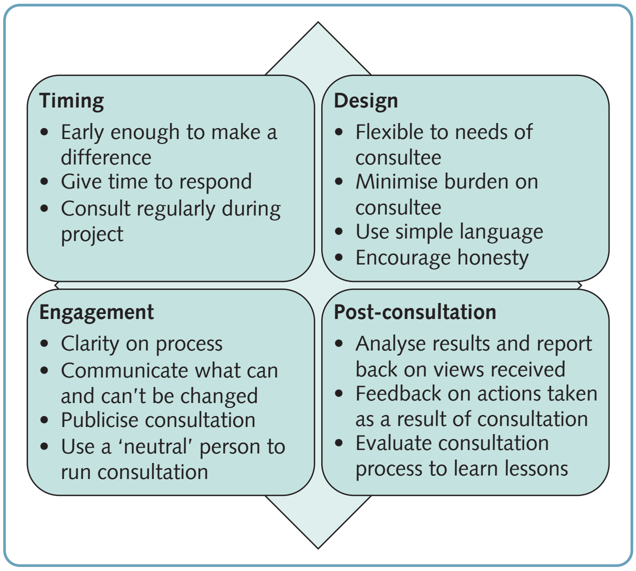


Figure 15.7 **Ensuring effective consultation with project stakeholders**

In seeking to understand these questions, stakeholder consultation becomes a critical activity. Consultation can provide valuable insights and experiences, can improve the legitimacy and buy-in for decisions, can help support relationships with key stakeholders, and can be critical in reducing potential opposition to the project. Figure 15.7 illustrates a number of key considerations for effective consultation with project stakeholders, considering timing, design, engagement and post-consultation.

Managing project stakeholders

Having identified stakeholders and understood their different perspectives on the project, the next step is to decide how best to manage the different stakeholders. One method is a differentiated approach based on stakeholder power and influence. Stakeholders who have the power to exercise a major influence over the project should never be ignored. At the very least, the nature of their interest, and their motivation, should be well understood. But not all stakeholders who have the power to exercise influence over a project will be interested in doing so, and not everyone who is interested in the project has the power to influence it. The power–interest grid, shown in Figure 15.8, classifies stakeholders simply in terms of these two dimensions. Although there will be graduations between them, the two dimensions are useful in providing an indication of how stakeholders can be managed in terms of these four categories.

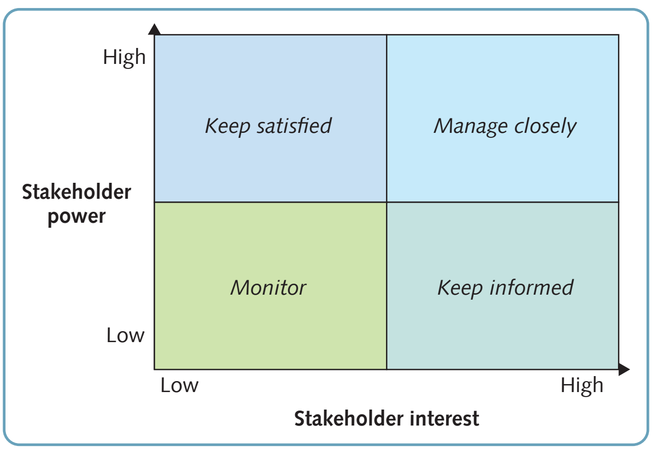


Figure 15.8 **Managing project stakeholders based on power and interest**

High-power and interested groups must be fully engaged, with the greatest efforts made to satisfy them. High-power and less-interested groups require enough effort to keep them satisfied, but not so much that they become bored or irritated with the message. Low-power and interested groups need to be kept adequately informed, with checks to ensure that no major issues are arising. These groups may be very helpful with the detail of the project. Low-power and less-interested groups need monitoring, though without excessive communication.

Case example

Berlin Brandenburg Airport opens at last[3]

Originally intended to replace the German capital's three ageing airports, the Berlin Brandenburg Airport is a major source of embarrassment for a country renowned for delivering things on time and on budget. The infrastructure project, one of the country's largest for decades, was due to open in 2011 with an expected 27 million passengers per year. However, after several major delays and failures, the airport finally received its licence to operate in May 2020 and didn't start operating until late 2020. It's not only the timing of the project that went badly wrong either – the initial budget of €2.83 billion more than doubled to over €7 billion. In fact, it is estimated that it cost around €20 million per month just to run the empty terminal building prior to its opening, plus €13 million per month in lost rental income.

What went so badly wrong? First, the growing popularity of Berlin as a destination meant that the original demand forecasts, made in 2006, were too low. Revised estimates (pre-COVID-19) indicated that the airport now needed to be able to handle somewhere between 30–35 million passengers per year (Dubai International, the world's busiest airport, handles around 90 million passengers). This led to investment in additional terminal space (especially around security, check-in and luggage reclaim), calls for a third runway to be developed, and a request by Hartmut Mehdorn, the industrial troubleshooter brought in to save the failing project, to keep one of the old airports, intended for closure, open to deal with the excess demand. Other problems included the airport's fire safety system – an innovative solution that, in the event of fire, pumps smoke under the terminal building rather than through the roof, but which failed to gain regulatory approval for a number of years. Over 1,000 automatic doors in the terminal building had to be re-engineered to ensure that they would close properly in the event of a fire. Other additional costs included additional parking, check-in counters and aircraft gates, rebuilding the airport's entrance hall, extending luggage facilities, and other cost overruns caused by cracking concrete in car parks, re-fitting of pipes and cables, missing conveyor belts, and problems with fire safety walls between the train station and terminal building. To add to these problems, the decision by airport bosses to cancel the contracts of the original consortium of architects and engineering firms led to significant rework in planning, as many of the documents and construction expertise became inaccessible.

Then, when it looked as if the project's troubles might finally be over, the coronavirus (COVID-19) pandemic caused a sudden and extreme drop in air travel. As it entered a new decade, Brandenburg Airport faced a new battle to regain old business and capture new clients in a sector that was in crisis. The need for increased capacity, a critical issue during construction, was now more questionable as Berlin's two main airports, Tegel and Schönefeld, experienced 65 per cent reductions in passenger numbers and 17 per cent reductions in cargo at the height of pandemic. While such dramatic reductions in flight statistics were evident in airports around the world, Brandenburg Airport faced a particularly difficult challenge given its less-established profile and its troubled history. It faced the strong prospect of significant reductions in business from critical intended customers in the low-cost carrier segment, such as easyJet and Ryanair. In addition, one of its large network carriers, Lufthansa, expected to focus its revival efforts on its main hubs in Munich and Frankfurt, instead of its secondary hub in Berlin.

peter jesche/Shutterstock

15.4 Diagnostic question: *Has the project been effectively defined?*

Before starting the complex task of planning and executing a project, it is essential to be clear about exactly what the project is – its definition. Three different elements define a project:

1. ***its objectives*** – the end state that the project is trying to achieve;
2. ***its scope*** – the exact range of the responsibilities taken on by the project;
3. ***its strategy*** – how the project is going to meet its objectives.

Project objectives

Objectives help to provide a definition of the project's end point and can therefore be used to monitor progress and identify when success has been achieved. Good objectives are those that are clear, measurable and, preferably, quantifiable. Clarifying objectives involves breaking down project objectives into three categories – the purpose, the end results and the success criteria. They can be judged in terms of the five performance objectives – quality, speed, dependability, flexibility and cost. However, flexibility is regarded as a 'given' in most projects that, by definition, are to some extent one-offs, and speed and dependability are compressed into one composite objective – 'time'. This results in what is known as the 'iron triangle' of project management – quality, time and cost. Although one objective might be particularly important, the other objectives can never be totally forgotten. As projects seek improved performance in one dimension (for example reducing the time of the project), we typically see reduced performance in one or both of the remaining performance dimensions (for example, reducing specification of the project scope and/or increasing the project budget).

OPERATIONS PRINCIPLE
Different projects will place different levels of emphasis on cost, time and quality objectives.

Project scope

Project scoping is a boundary-setting exercise that attempts to define the dividing line between what each part of the project will and won't do. Defining scope is particularly important when part of a project is being outsourced. A supplier's scope of supply will identify the legal boundaries within which the work must be done. Sometimes the scope of the project is articulated in a formal 'project specification'. This is the written, pictorial and graphical information used to define the output, and the accompanying terms and conditions. The project scope will also outline limits or exclusions to the project. This is critical, because perceptions of project success or failure often originate from the extent to which deliverables, limits and exclusions have been clearly stated and understood by all parties during the scoping phase.

Project strategy

The third part of a project's definition is the project strategy, which defines, in a general rather than a specific way, how the project is going to meets its objectives. It does this in two ways: by defining the phases of the project, and by setting milestones and/or 'stagegates'. Milestones are important events during the project's life and often trigger payments to contractors. Stagegates are the decision points that allow the project to move on to its next phase. A stagegate often launches further activities and therefore commits the project to additional costs, etc. During project definition, the actual dates for each milestone or stagegate are not necessarily determined. However, identifying the significant project milestones and stagegates is very helpful in supporting discussions with key stakeholders and in clarifying boundaries between project phases.

Case example

The risk of changing project scope – the Vasa project[4]

(This example was written and kindly supplied by Professor Mattia Bianchi, Stockholm School of Economics)

Project specification changes, alongside poor communication and simple bad luck, have always had the ability to bring down even the most high-profile projects. In 1628, the Vasa, the most magnificent warship ever built for the Royal Swedish Navy, was launched in front of an excited crowd. It had sailed less than a few thousand metres during its maiden voyage in the waters of the Stockholm harbour when suddenly, after a gun salute was shot in celebration, the Vasa heeled over. As water gushed in through the gun ports, the ship vanished beneath the surface killing 53 of the 150 passengers. Shocked officials were left questioning how such a disaster could happen.

Yet, as a project, the story of the Vasa displayed many of the signs of potential failure. When her construction began in 1625, the Vasa was designed as a small traditional warship, similar to many others previously built by the experienced shipbuilder Henrik Hybertsson. Soon after, the Swedish King, Gustav II Adolphus, at that time fighting the Polish Navy in the Baltic Sea, started ordering a series of changes to the shape and the size of the warship, making its design much longer and bigger than originally envisaged. In addition, his spies informed the King that the Danes had started building warships with two gun decks, instead of the customary one. This would give them a great advantage in terms of superior firepower from a longer distance. From the battlefront, the King ordered the addition of a second gun deck to the Vasa. The message caused consternation when it reached the shipbuilder several months later, but they attempted to comply with the change even though it caused wasteful reworking and complex patching up, as no one had ever seen or built such a revolutionary design before. Yet more pressure was put on the project when a major storm destroyed ten of the King's ships, making the commissioning of the Vasa even more urgent. Then, as a final piece of bad luck (especially for him) the shipbuilder, Hybertsson, died. Nevertheless, just before the ship's completion, a Navy representative, Admiral Fleming, conducted a stability test to assess the seaworthiness of the ship. Notwithstanding the strong signals of instability, the Vasa was launched on its maiden voyage – with disastrous results for the King, for the Swedish Navy and the project. The example highlights the major risks of interventions in projects to (radically) change their scope. In this case, not only was Vasa's specification changed, but the schedule was compressed, creating a high risk of project failure.

15.5 Diagnostic question: *Has the project been adequately planned?*

All projects, even the smallest, need some degree of planning. The planning process fulfils four distinct purposes:

1. It determines the cost and duration of the project. This enables major decisions to be made, such as the decision whether or not to go ahead with the project in the first place.
2. It determines the level of resources that will be needed.
3. It helps to allocate work and monitor progress. Planning must include the identification of who is responsible for what.
4. It helps to assess the impact of any changes to the project.

OPERATIONS PRINCIPLE
Project planning is essential for all types of projects, but especially those with higher levels of innovation, pace or complexity.

Planning is often repeated several times during the project's life as circumstances change. As discussed earlier, projects can and should be differentiated based on their characteristics – in our case, we have examined the level of innovation, pace and complexity. When managing particularly difficult projects, it is therefore a normal occurrence to revise plans as the project progresses. The process of project planning involves five steps, as shown in Figure 15.9.

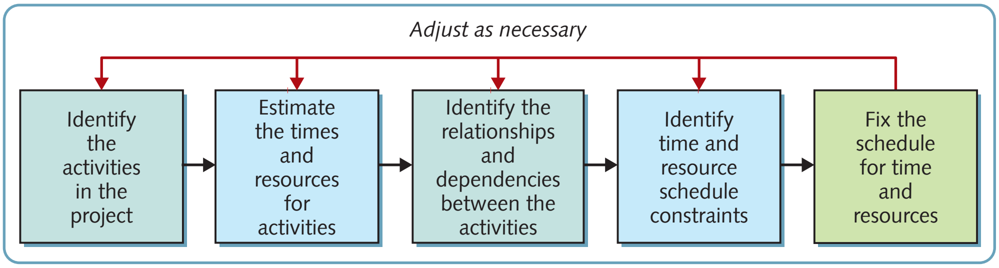


Figure 15.9 **Stages in the project planning process**

Identify activities – the work breakdown structure

Some projects are too complex to be planned and controlled effectively unless they are first broken down into manageable portions. This is achieved by structuring the project into a 'family tree' that specifies the major tasks or sub-projects. These in turn are divided up into smaller tasks until a defined, manageable series of tasks, called work packages, is arrived at. Each work package can be allocated its own objectives in terms of time, cost and quality, and can be assigned specific responsibility for its delivery. Typically, work packages do not exceed 10 days (though in practice they often do), should be independent from each other, should belong to one sub-deliverable, and should be monitored constantly. The output from this is called the work breakdown structure (WBS). The WBS brings clarity to the project planning process and provides a framework for building up information for reporting purposes.

For example, Figure 15.10 shows the work breakdown structure for a project to design a web-based interface for a new sales knowledge management system that is being installed in an insurance company. The project is a cooperation between the company's IT systems department and its sales organisation. Three types of activity will be necessary to complete the project – training, installation and testing. Each of these categories is further broken down into specific activities.

Estimate times and resources

The next stage in project planning is to identify the time and resource requirements of the work packages. Estimates are rarely perfect, but they can be made with some idea of how accurate they might be. Table 15.2 includes time and resource estimates for the sales system interface design project.

There are two approaches typically taken in estimating time or resource needs for a project. Top-down estimates look at the project as a whole and typically use an *analogy approach* (for example, estimating the time of an NPSD project based on similar previous projects), a *ratio approach* (for example, estimating the cost of building a new house using

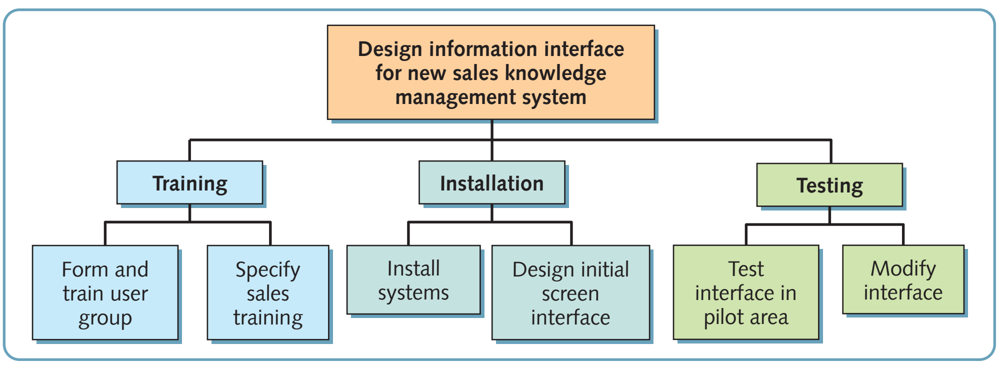


Figure 15.10 **Work breakdown structure for a project to design an information interface for a new sales knowledge management system in an insurance company**

Table 15.2 Time, resources and relationships for the sales system interface design project

Code	*Activity*	*Immediate predecessor(s)*	*Duration (days)*	*Resources (developers)*
a	Form and train user group	none	10	3
b	Install systems	none	17	5
c	Specify sales training	a	5	2
d	Design initial screen interface	a	5	3
e	Test interface in pilot area	b, d	25	2
f	Modify interface	c, e	15	3

a cost-per-square-metre calculation) or a *consensus approach* (where a group with experience of similar projects discusses the project to form a best estimate). Top-down methods are most commonly adopted when very precise estimations are not required or not possible (for example, for highly uncertain projects). Bottom-up approaches to estimation focus on breaking down the project into smaller parts and then estimating the time or resource requirements for each of these parts. With bottom-up estimation, project managers are typically relying on those who will actually be doing the work to come up with an accurate estimate.

Identify the relationships and dependencies between the activities

The third stage of project planning is to understand the interactions between different project work packages. All the work packages that are identified as comprising a project will have some relationship with one another. Some activities need to be executed in a particular order while other activities are independent of one another and can be carried out in parallel. In the case of the sales system interface design, Table 15.2 provides the basic information that enables the relationships between activities in the project to be established. It does this by identifying the immediate predecessor (or predecessors) for each activity. So, for example, activities **a** and **b** can be started without any of the other activities being completed. Activity **c** cannot begin until activity **a** has been completed, nor can activity **d**. Activity **e** can start only when both activities **b** and **d** have been completed, and activity **f** can start only when activities **c** and **e** have been completed.

Planning tools

Project planning is greatly aided by the use of techniques that help to handle time, resource and relationship complexity. Figure 15.11 shows a Gantt chart (introduced in Chapter 10) for the

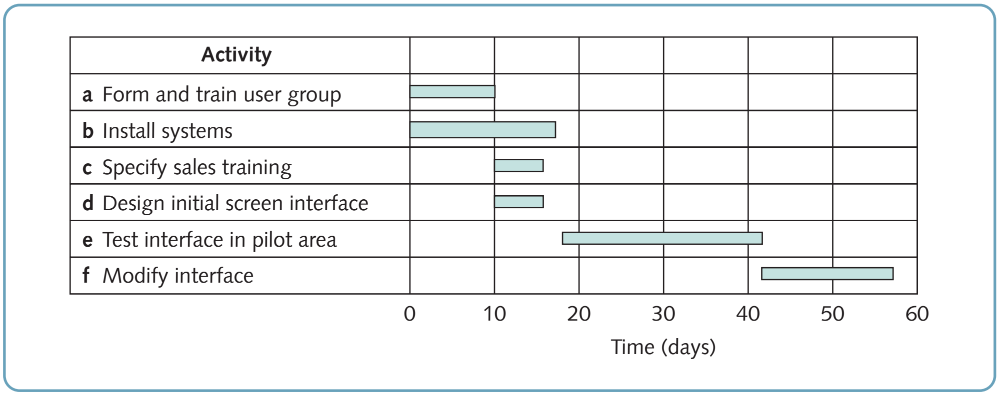


Figure 15.11 Gantt chart for the project to design an information interface for a new sales knowledge management system in an insurance company

activities that form the sales system interface project. The bars indicate the start, duration and finish time for each activity. Gantt charts are the simplest way to exhibit an overall project plan, because they have excellent visual impact and are easy to understand. They are also useful for communicating project plans and status to senior managers, as well as for day-to-day project control.

Network analysis

As project complexity increases, it becomes more necessary to identify clearly the relationships between activities and show the logical sequence in which activities must take place. This is most commonly done by using the *critical path method* (CPM) to clarify the relationships between activities diagrammatically. Though there are alternative methods of carrying out critical path analysis, by far the most common, and also the one used in most project management software packages, is the 'activity on node' (AoN) method. Figure 15.12 shows this for the sales team interface design project.

In the AoN representation, activities are drawn as boxes, and arrows are used to define the relationships between them. In the centre of each box is the description of the activity. Above the description are the duration (D) of the activity (or work package), the earliest start time (EST) and earliest finish time (EFT). Below the description are the latest start time (LST), the latest finish time (LFT), and the 'float' (F) (the number of extra days that the activity could take without slowing down the overall project). In this case, activity chains **a** – **c** – **f**, and **a** – **d** – **e** – **f**, and **b** – **e** – **f**, must all be completed before the project can be considered as finished. The longest (in duration) of these chains of activities is called the 'critical path' because it represents the shortest time in which the project can be finished, and therefore dictates the project timing. In this case **b** – **e** – **f** is the longest path and the earliest project finish is 57 days.

Activities that lie on the critical path will have the same earliest and latest start times and earliest and latest finish times. Non-critical activities, however, have some flexibility as to when they start and finish. This flexibility is quantified into a figure that is known either as 'float' or 'slack'. So, activity **c**, for example, is of only 5 days duration and it can start any time after day 10 (when activity **a** is completed) and must finish any time before day 42 (when activities **a**, **b**, **c** and **d** are completed). Its 'float' is therefore 42 − 10 − 5 = 27 days (i.e. latest finish time minus earliest start time minus activity duration).

Identify time and resource schedule constraints

The logic that governs project relationships, as shown in the critical path analysis, is primarily derived from the technical details, but the availability of resources may also impose its own

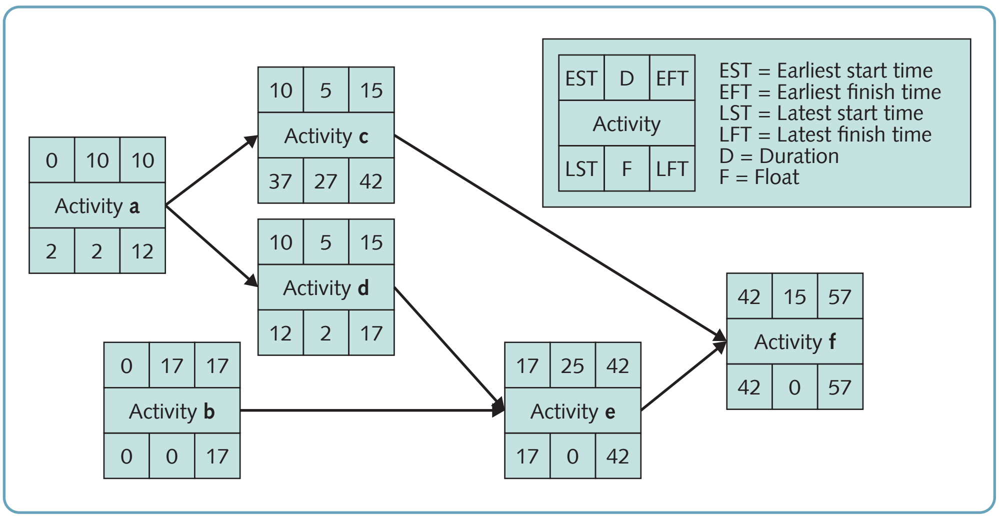


Figure 15.12 Critical path analysis (activity-on-node method) for the project to design an information interface for a new sales knowledge management system in an insurance company

constraints, which can materially affect the relationships between activities. For example, specialist staff may not have the available time to carry out two tasks simultaneously even if the critical analysis has identified that two activities can *technically* run in parallel. This often has the effect of highlighting the need for more detailed replanning.

Return to the sales system interface design project. Figure 15.13 shows the resource profile under two different assumptions. The critical path activities (**b** – **e** – **f**) form the initial basis of the project's resource profile. These activities have no float and can only take place as shown. However, activities **a**, **c** and **d** are not on the critical path, so project managers have some flexibility as to when these activities occur, and therefore when the resources associated with these activities will be required. From Figure 15.13, if one schedules all activities to start as soon as possible, the resource profile peaks between days 10 and 15, when 10 IT development staff are required. However, if the project managers exploit the float for activity **c** and delay its start until after activity **b** has been completed (day 17), the number of IT developers required by the project does not exceed 8. In this way, float can be used to smooth resource requirements or make the project fit resource constraints. However, it does impose further resource-constrained logic on the relationship between the activities. So, for example, in this project, moving activity **c** as shown in Figure 15.13 results in a further constraint of not starting activity **c** until activity **b** has been completed.

Fix the schedule for time and resources

Project planners should ideally have a number of alternatives to choose from. The one which best fits project objectives can then be chosen or developed. While it can be challenging to examine several alternative schedules, especially in very large or very uncertain projects, computer-based software packages such as Bitrix24, Trello, Asana, MS Project and Producteev make critical path optimisation more feasible. The rather tedious computation necessary in network planning can be relatively easily performed by project planning models. All they need are the basic relationships between activities together with timing and resource requirements for each activity. Earliest and latest event times, float and other characteristics of a network can then be presented, often in the form of a Gantt chart. More significantly, the speed of computation allows for frequent updates to the project plans. Similarly, if updated information is both accurate and frequent, such software can provide effective project control data.

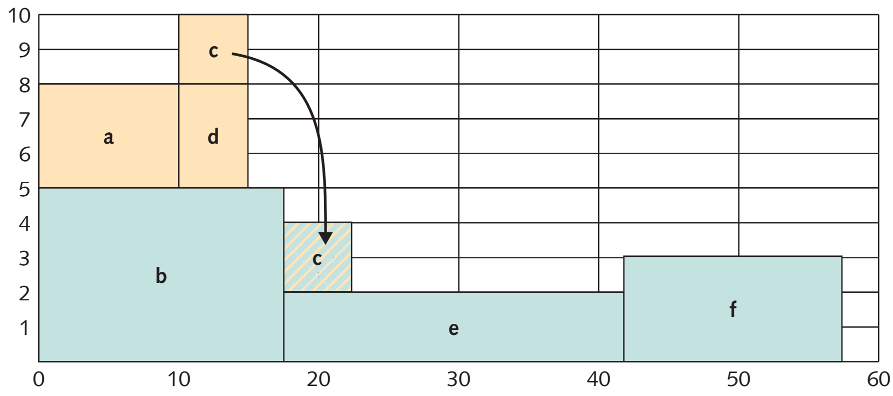
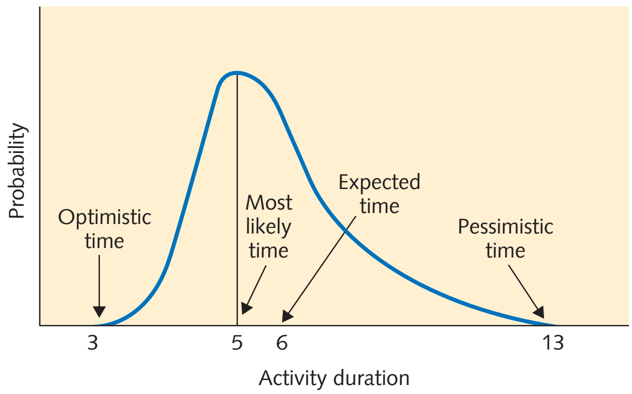


Figure 15.13 Resource profiles for the sales knowledge system interface design, assuming that all activities are started as soon as possible, and assuming that the float in activity c is used to smooth the resource profile

Programme evaluation and review technique (PERT)

While it is beyond the scope of this text to enter into much more detail of the various ways that critical path analysis can be made more sophisticated, programme evaluation and review technique (PERT) is worth noting given its popularity among practising project managers. PERT, as it is universally known, originated in the planning and control of major defence projects in the US Navy, with its most spectacular gains in the highly uncertain environment of space and defence projects. The technique recognises that activity durations and costs in projects are not deterministic (fixed), but instead a probability curve can be used to describe the estimate. The natural tendency of some people is to produce optimistic estimates (optimism bias is one of a number of cognitive biases discussed in the critical commentary at the end of this chapter), but these will have a relatively low probability of being correct because they represent the time or cost if everything goes very well. Most likely estimates have the highest probability of proving correct. Finally, pessimistic estimates assume that almost everything that could go wrong does go wrong. In practice, this is usually a positively skewed distribution (as shown in Figure 15.14 for a time estimate), so the expected time (or cost) will not be the same as the most likely time (or cost). More uncertainty increases the range of the distribution. Table 15.3 shows a PERT analysis for the time of different activities in the sales system interface design project that we have discussed previously. The same approach can also be used to develop more realistic project cost

> **OPERATIONS PRINCIPLE**
> ***Probabilistic activity-time estimates facilitate the assessment of a project being completed on time.***



$$\text{Expected activity time } (t_e) = \frac{t_o + 4t_1 + t_p}{6}$$

$$\text{Variance} = V = \frac{(t_p - t_e)^2}{36}$$

where

t_o = optimistic activity time

t_l = most likely activity time

t_p = pessimistic activity time

Figure 15.14 Using probabilistic estimates in project planning

Table 15.3 PERT analysis for the sales system interface design project

Code	*Activity*	*Optimistic estimate*	*Most likely estimate*	*Pessimistic estimate*	*Expected time*	*Variance*
a	Form and train user group	8	10	14	10.33	1
b	Install systems	10	17	25	17.17	6.25
c	Specify sales training	4	5	6	5	0.11
d	Design initial screen interface	5	5	5	5	0
e	Test interface in pilot area	22	25	27	24.83	0.69
f	Modify interface	12	15	25	16.17	4.69

Case example

Ocado's robotics projects[5]

Ocado, the online grocery retailer, remains at the forefront of technology to support its growing operations. While the use of automated warehouse systems is by no means a new phenomenon, the pace of technology adoption has risen sharply in recent years as a result of both rising labour costs and the availability of better and more cost-effective technologies. Ocado has a number of projects seeking to leverage new technological opportunities. One is the development of advanced packing robots, capable of handling heavy or hazardous products (to avoid worker injuries) as well as delicate objects, such as fruits, vegetables, salads and eggs. Another recent technology project for Ocado is the development of a humanoid assistant (think of C-3PO from the *Star Wars* movies, but with wheels instead of legs!) aimed at supporting engineers in the maintenance of its product handling systems. In partnership with the Karlsruhe Institute of Technology in Germany, Ecole Polytechnique Fédérale de Lausanne in Switzerland, University College London in the UK, and Sapienza University in Italy, these robots offer a pair of 'second hands' to engineers, moving tools and materials and handing them to their human partners as needed. They are also capable of interrupting human actions to offer advice on alternative solutions to common problems. According to Ocado, the aim is to create a fluid and natural interaction between robot and technician within its operations. These two examples point to the continually changing nature of the workplace, as technology is increasingly integrated within many tasks. It also highlights the value (and challenges) of bringing together different areas of expertise from geographically dispersed partners in order to deliver project success.

Sorn340 Studio Images/Shutterstock

estimates. In this case, the sum of the expected times for each of the activities on the critical path (**b** – **e** – **f**) is 58.17 days (as opposed to the original 57 days based on 'most likely' estimates for these activities) and the sum of the variances of these three activities is 11.63 days. From this, project managers can calculate the probability of the project overrunning by different amounts of time.

15.6 Diagnostic question: *Is the project effectively controlled and learned from?*

Understanding the project environment and its stakeholders, project definition and project planning all take place largely before the actual project begins. By contrast, project control and learning deals with the management activities that take place during the execution of the project and after it ends. It involves the following key activities:

- *Monitoring* the project in order to check on its progress.
- *Assessing the performance* of the project by comparing monitored observations of the project with the project plan.
- *Intervening* in the project in order to make the changes that will bring it back to plan.
- *Learning* from the project in order to improve the performance of subsequent projects.

Project monitoring

Project managers have first to decide what they should be looking for as the project progresses. Common measures include current expenditure to date, supplier price changes, amount of overtime authorised, technical changes to project, inspection failures, number and length of delays, activities not started on time, missed milestones, etc. Some of these monitored measures affect mainly cost, some mainly time. However, when something affects the quality of the project, there are also time and cost implications. This is because quality problems in project planning and control usually have to be solved in a limited amount of time and within a budget.

Assessing project performance

A typical planned cost profile of a project through its life is shown in Figure 15.15. At the beginning of a project, some activities can be started but most activities will be dependent on others finishing. Eventually, only a few activities will remain to be completed. This pattern of a slow start followed by a faster pace with an eventual tail-off of activity holds true for almost all projects, which is why the rate of total expenditure follows an S-shaped pattern. It is against this curve that actual costs can be compared in order to check whether the project's costs are being incurred to plan. Figure 15.15 shows the planned and actual cost figures compared in this way. It shows that the project is incurring costs, on a cumulative basis, ahead of what was planned.

Earned value analysis

Earned value analysis (EVA) is a technique that allows for various comparisons to be made between expected costs and schedules, and the actual performance of a project. Table 15.4 illustrates an EVA for a simple project. This technique is not only useful for determining how well a project is progressing in terms of tasks completed and costs incurred, it also helps in re-evaluating original budgets and time schedules. In this case (as of week 6 when the EVA was carried out), the project is running 11.4 per cent over budget and 19.5 per cent behind schedule. If things continue in this way for the remainder of the project, it is likely to cost around €117,381 (as opposed to the original budget of €104,000) and be delivered in 12.4 weeks (as opposed to the planned 10 weeks).

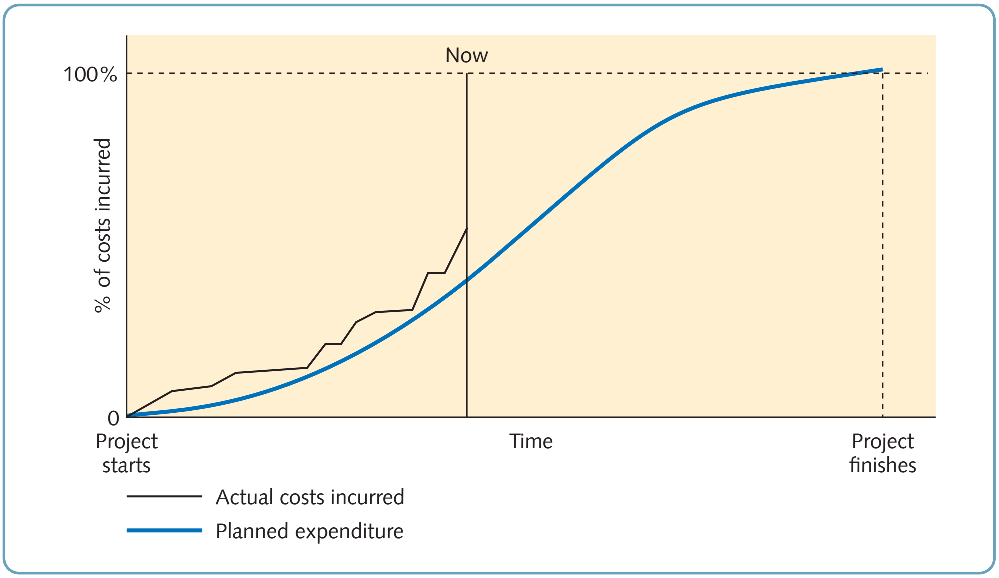


Figure 15.15 **Comparing planned and actual expenditure**

Table 15.4 Earned value analysis (EVA) for a simple project

Activity	Planned time	Planned budget
1	1 week	€5,500
2	1 week	€8,750
3	1 week	€6,250
4	1 week	€11,000
5	1 week	€15,000
6	1 week	€11,250
7	1 week	€13,750
8	1 week	€9,000
9	1 week	€14,000
10	1 week	€9,500
TOTAL	**10 weeks**	**€104,000**

PROJECT REVIEW – END WEEK 6:
* Work complete: Activities 1–5
* Actual cost (AC) at time of review = €52,500
* Planned value (PV) = sum of weeks 1–6 = €57,750
* Earned value (EV) = sum of activities completed = €46,500

COST review:
* Cost variance (CV) = EV − AC = €46,500 − €52,500 = (€6,000) (Negative CV, overspend)
* Cost performance index (CPI) = EV / AC = €46,500 / €52,500 = 0.886 (CPI < 1 overspend)
* Estimate at completion (EAC) = Budget at completion (BAC) / CPI = €104,000 / 0.886 = €117,381

SCHEDULE review:
* Schedule variance (SV) = EV − PV = €46,500 − €57,750 = (€11,250) (Negative SV, behind schedule)
* Schedule performance index (SPI) = EV / PV = €46,500 / €57,750 = 0.805 (SPI < 1, behind schedule)
* Estimated time to complete (ETC) = Original time estimate / SPI = 10 / 0.805 = 12.4 weeks

Intervening to change the project

If the project is obviously out of control in the sense that its costs, quality levels or times are significantly different from those planned, then some kind of intervention is almost certainly required. Given the interconnected nature of projects, a change to one part of the project will have knock-on effects elsewhere. Therefore, interventions often require wide consultation. One common form of intervention is to 'crash' activities. Crashing is the process of reducing time spans for critical path activities so that the project can be completed in less time. Usually, crashing activities incurs extra costs in terms of overtime working, additional resources or sub-contracting work.

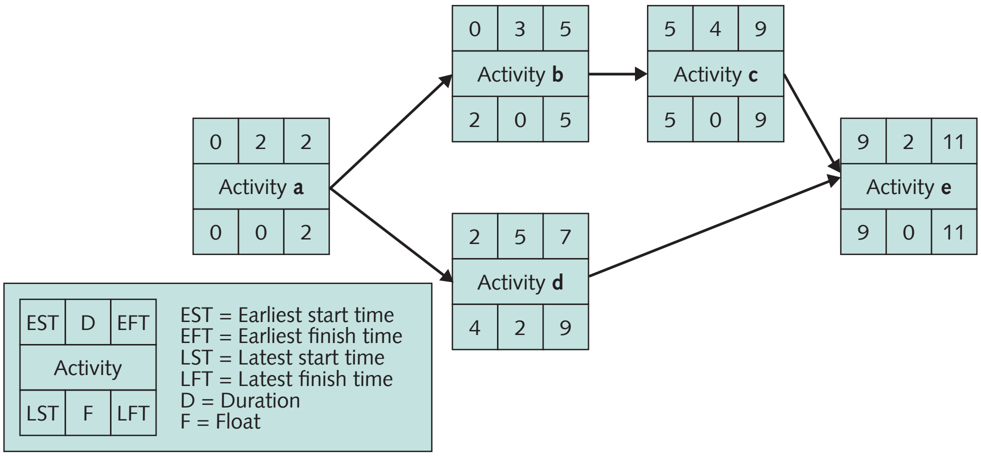
Figure 15.16 shows an example of crashing a simple project network. For each activity, the duration and normal cost are specified, together with the (reduced) duration and (increased) cost of crashing them. If the total project time is to be reduced, one of the activities on the critical path (**a**, **b**, **c**, **e**) must be crashed. Not all activities are capable of being crashed – in this case, activity **e**. To decide which activity to crash, the 'cost slope' of each is calculated. This is the cost per time period of reducing durations. The most cost-effective way of shortening the whole project then is to crash the activity on the critical path with the lowest cost slope. This is activity **a**, the crashing of which will cost an extra €2,000 and will shorten the project by one week. After this, activity **c** can be crashed, saving a further two weeks and costing an extra €5,000. At this point all the activities have become critical and further time savings can only be achieved by crashing two activities in parallel.

The shape of the time–cost curve in Figure 15.16 is entirely typical. Initial savings come relatively inexpensively if the activities with the lowest cost slope are chosen. Later in the crashing sequence, the more expensive activities need to be crashed and eventually two or more paths become jointly critical. Inevitably by that point, savings in time can only come from crashing two or more activities on parallel paths.

OPERATIONS PRINCIPLE
Only accelerating activities on the critical path(s) will accelerate the whole project.

Managing project learning

The activity of project management doesn't stop when a project comes to an end – managing the process of project learning is critical to future project performance. Yet, within the majority of projects, there remains very little formalised learning. This can partly be explained by the key



Activity	*Normal*		*Crash*		*Cost slope (€ 000/week)*
	Cost (€ 000)	*Time (weeks)*	*Cost (€ 000)*	*Time (weeks)*	
a	6	2	8	1	2
b	5	3	8	2	3
c	10	4	15	2	2.5
d	5	5	9	4	4
e	7	2	Not possible		–

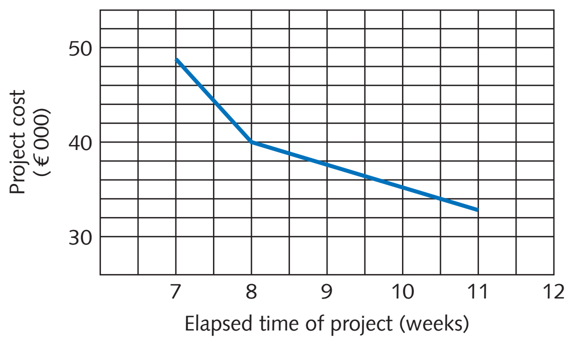


Figure 15.16 Crashing activities to shorten project time becomes progressively more expensive

performance objectives for individuals involved in projects – typically focused on the success of an individual project (in terms of quality, time, cost) as opposed to longer-term learning effects or the development of organisational capabilities. As a result, when the project ends, there may be little incentive for stakeholders to spend time reviewing aspects of the project's execution that could have been improved. In addition, when things go wrong, those involved often prefer to move on rather than go back and examine mistakes made. However, not incorporating learning as a key part of all projects is a huge missed opportunity. Remember, most projects fail in some way: failure to fully understand the project environment and key stakeholders; failure to accurately define the project; failure in time and resource estimation, scheduling or constraint identification; failure in performance monitoring and intervention during project delivery; or failure to manage the interests of different stakeholders involved in a project. Conversely, when mistakes *are* actively learned from, they can provide significant performance benefits over time, as illustrated in Figure 15.17. Where organisations have limited formalised learning mechanisms

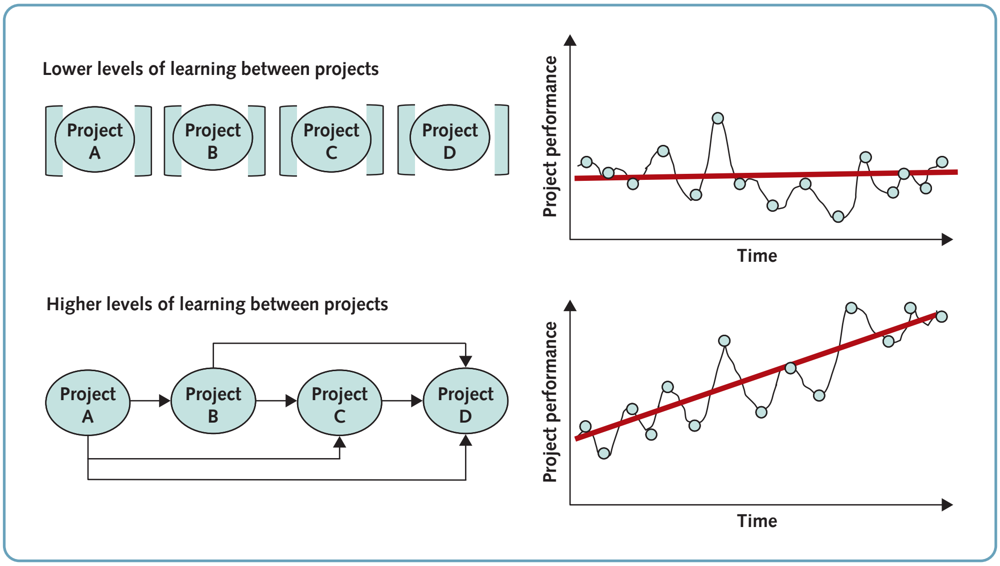


Figure 15.17 **Improving project performance over time through learning between projects**

as part of their project processes, there will be little change in the underlying *average* performance of their projects over time. But where organisations place a greater emphasis on formalised learning from one project to another, there is typically an *upwards trend* in project performance (despite variance in project performance remaining).

Critical commentary

- While the importance of understanding and managing stakeholders is increasingly acknowledged in project management, there remain different perspectives on the fundamental role of both stakeholders and stakeholder management. Some argue that stakeholder satisfaction should not be seen as a goal in and of itself, but rather stakeholders should be managed only when they have a direct impact on the project's outcomes. Others take a broader view and see stakeholders as having ownership of the project. As such, maximising their welfare, alongside stockholders, becomes a key priority for project management. Finally, some question the very idea of stakeholder management, arguing that it complicates managerial practice and fundamentally challenges the idea of the corporate objective function. Further, they note that adopting a stakeholder perspective provides an easy excuse for poor project management decisions, as any decision can retrospectively be presented as an attempt to respond to stakeholder needs.
- When exploring project definition, we noted the trade-offs between quality, time and cost – the so-called 'iron triangle'. When adopting this perspective, project managers must balance improvements in one area of project performance against likely deterioration in one or both of the remaining elements. So, to speed up a project, we can expect increased costs (i.e. cost performance worsens) and possible reductions in deliverables (i.e. quality performance worsens). Although many find the iron triangle perspective useful in making trade-offs explicit during the scoping of projects, others argue that it creates a constrained mindset. When problems present themselves they are not addressed creatively, but instead simply lead to requests for additional budget or time. In addition, some project scopes are simply unfeasible regardless of expanding time and budget envelopes. In other cases, adding budget and resources can sometimes actually slow activities given the effects of increased coordination and communication complexity.

- When project managers talk of 'estimates', they are really talking about guessing. By definition, planning a project happens in advance of the project itself. Therefore, no one really knows how long each activity will take or what it will cost. Of course, some kind of guess is needed for planning purposes. However, some project managers believe that too much faith is put in time and cost estimates. The really important considerations, they claim, are how long something could take without delaying the whole project and how much something could cost without it harming the project's viability. Also, if a single most likely estimate is unreliable, then using three, as one does for probabilistic estimates, is merely over-analysing what is highly dubious data in the first place. It is also important for project managers to be aware of the likely biases that they and their team may be affected by when developing estimates for time and costs. Earlier, when discussing PERT, we noted the effect of optimism bias (i.e. the tendency to overestimate the likelihood of positive events occurring and underestimate the likelihood of negative events occurring) in estimates. In addition, project planners may be affected by, for example:
  - *Anchoring bias*: An over-reliance on an initial piece of information that 'anchors' subsequent judgements (for example, the anchoring effect of an initial cost estimate for a project).
  - *Bandwagon effects*: A form of groupthink where individuals believe something because others do (for example, making a similar time estimate to those of other members of the project team).
  - *Recency bias*: An over-reliance on more recent forms of information, relative to older forms of information (for example, estimates of project supplier risk dominated by experiences from the last completed project).
  - *Confirmation bias*: The tendency to search for and select pieces of information that confirm, rather than refute, a given position (for example, purposely selecting examples of activity time completions from previous projects that support a time estimate being made for a new project).

SUMMARY CHECKLIST

- ☐ Do managers understand what projects are, and their innate characteristics?
- ☐ Do managers understand the implications of differences in levels of project innovation, pace and complexity?
- ☐ Is the activity of project management understood and are managers clear on the key responsibilities of a project manager?
- ☐ For this particular project, does the project manager have skills appropriate for the project's intrinsic degree of difficulty?
- ☐ Do managers understand the key characteristics of the project environment in terms of internal, business, economic–political and geo-social dimensions?
- ☐ Do managers understand the importance of identifying and managing project stakeholders?
- ☐ Do managers have methods of stakeholder mapping, consultation and management?
- ☐ Have the objectives of the project been defined, particularly in terms of the relative importance of cost, time and quality?
- ☐ Has the scope of the project been defined, including technical specifications, limits and exclusions?
- ☐ Has the overall strategy of the project been defined in terms of its overall approach, its significant milestones and any decision gateways (stagegates) that may occur in the project?
- ☐ Is sufficient effort being put into the project planning process?
- ☐ Have all activities been identified and expressed in the form of a work breakdown structure?

- ☐ Have all activity times and resources been estimated using the best possible information?
- ☐ Is there sufficient confidence in the time and resource estimates to make planning meaningful?
- ☐ Have the relationships and dependencies between activities been identified and summarised in the form of a simple network diagram?
- ☐ Have project planning tools, such as critical path analysis or PERT, been considered for the project?
- ☐ Have potential resource and time schedule constraints been built into the project plan?
- ☐ Are there mechanisms in place to monitor the progress of the project?
- ☐ Have mechanisms for intervening in the project to bring it back to plan been put in place?
- ☐ What methods of project learning have been designed into the project process to increase the likelihood of improved performance for subsequent projects?

Case study

Kloud BV and Sakura Bank K.K.

(This case was co-authored with Nigel Spinks, Henley Business School, University of Reading)

'*Well that's the bad news!*', said Tao, the Managing Director of Kloud BV, a consulting and executive development firm headquartered in Amsterdam, specialising in operations and supply chain improvement. '*The good news is that Chao should be out of hospital in a couple of weeks. It may take a few months before he's fully fit, but it all looks very promising*.' Maria was pleased to hear that things were looking more positive for Chao after his accident. She had only been at the company for six weeks, having taken up a role as a junior project manager, but had already grown to respect and like Chao.

'*But*', continued Tao, '*that does leave us in a tricky situation. As you know, Chao was in charge of the big project with Sakura Bank in Tokyo, which I'm going to look after until he's back at work. He was also just setting up a smaller project for them, training senior managers, which will run out of their facilities in Osaka. I appreciate you're pretty new here, but I'd like you to take on the project management for this one. Chao recommended you, so it seems you've made a very good impression!*' Maria was pleased to hear that Chao, her immediate boss, had a good impression of her. '*Well, I'm very happy to take this on Tao*', she said as she quickly looked through the draft proposal for the project that Chao had been developing for Sakura Bank just before his accident (see Figure 15.18).

As Maria read through the proposal, she got a clearer idea of what was needed, but she still had a number of questions. '*There's plenty of information for me here Tao. Still, what constraints do I need to be aware of?*' Tao picked up a notebook from the corner of his desk. '*Good question! I was chatting to Chao earlier today and he mentioned a few things. The client kick-off meeting takes place online next Monday – so that's week 1 on this project. Sakura have already said that, ideally, they'd like the residential programme to start in week 6. Do you think that's a realistic time frame? They're also pretty keen that pre-programme activities and the residential programme elements start on Mondays, and that Saturdays and Sundays are non-working days.*' Maria and Tao's discussion then moved to how best to resource the project. Within about ten minutes, they had identified most of the key players who would be involved:

- *Project sponsor* – Tao (attend online client kick-off meeting and final review with client; will review final report).
- *Project manager* – Maria to replace Chao (run client and kick-off meetings, sign-off trainer contracting and programme design, and do final report and client management).
- *Training lead* – Kavita in the Tokyo office (training needs analysis, identify trainers and detailed programme design, on-site lead for residential training).
- *Web design* – Li Wei in the Shanghai office (liaise with Kavita and Una).
- *Project support/admin* – Krister (distribute contracts, confirm travel/accommodation/meal bookings, etc. for the meetings/residential programme).
- *Training* – three external trainers (finalise names once training agenda is completed): prepare materials, support pre-programme online training; one trainer per week for residential training, supported by Kavita as training lead. Most likely three trainers: two days each to develop materials (extra two days internal time to review content, check for overlaps, etc.); three to four days each on online support for pre-programme activities and five days each on residential delivery.
- *Survey* – Una in the Shanghai office (design, distribution and analysis of final survey; discuss with Kavita).
- *Invoicing and budget support* – Ruben (track invoicing and do budget close for project).

Maria then turned her attention to an additional note that Chao had made on the key activities in the project, including their time estimates, predecessors and average daily costs (see Table 15.5).

Maria thought for a moment. She assumed that Chao had developed his time estimates for each activity based on normal costing but wasn't sure what options there might be to reduce the time of some of these activities. '*Tao, I don't suppose Chao made any notes on possible activity 'crashing' did he?*' After rummaging around his desk for what seemed like an age, Tao found a bright pink Post-it Note hiding under a collection of files. '*Phew, I was starting to think I'd lost this! So, it looks like the training needs analysis could be shortened to two days, but it'll increase the daily cost to €850; the programme design activity can be shortened from five days to four days, but daily costs will increase to €750; for a fixed fee of €4,000, we could get a single more experienced trainer to do the training material creation in four days; and website set-up could be done in three days, but daily costs will increase to €500 per day*.' Maria looked up from her notes. '*OK, that's good to know. Anything else?*' Tao took a sip of water. '*Well, I guess it's important to say that Sakura is an important new client. There's a lot of potential for growth if we can deliver this project and the one I'll be leading effectively! We've heard from a few other firms who've worked with them that they can be quite a challenging client – apparently, they often change their mind on specifications! Oh, and I nearly forgot, to ensure that any project is viable for Kloud, we*

Introduction

Kloud BV is delighted to submit a proposal for the design, development and delivery of a three-week executive training programme to equip your managers with the latest operations improvement insights from around the world. The initial programme will form the basis for future cohorts over the next three years.

Overall approach

The programme is to be delivered to a cohort of 12 managers selected by Sakura Bank K.K. from their global talent pool. Our approach is to combine online pre-programme activities to help participants prepare for the face-to-face component of the programme, which will be delivered by our experienced and dedicated team of operations improvement experts, using up-to-date and creative teaching strategies focused on practical application. Throughout the project, the training team will be supported by Kloud's own programme management team who will also be responsible for evaluating and reporting on programme outcomes to Sakura's senior management team.

Programme structure

Once participants have been selected for the programme, they will be asked to undertake a structured series of online activities to ensure that they all have a sound level of knowledge of basic operations improvement concepts. This will enable participants to get the best value out of the residential programme by avoiding loss of time on teaching topics that are known to most of the group. The three-week residential training programme will combine plenary knowledge-transfer sessions with individual and group discussions. The final programme is subject to agreement with Sakura Bank K.K., but will take place in three blocks, each lasting five days (as requested by the client in initial conversations). Evaluation is a core component of the programme in order to assess its impact for participants and Sakura Bank K.K. Post-programme evaluation will be undertaken to measure participant satisfaction and their knowledge of programme topics. Full reporting of findings will be provided to you.

Programme design

The programme design will leverage Kloud's expertise in operations improvement consulting and executive development delivered to our global client base. Nevertheless, we believe strongly in the need to tailor design and delivery to specific client needs. We therefore propose to work collaboratively with Sakura Bank K.K., including on-site training needs analysis, as part of the final programme design. Following delivery of this first programme, the design and materials will then be used for subsequent follow-on cohorts in other regions, as required by Sakura Bank K.K.

Kloud's delivery team

Kloud's delivery team will be led by Chao Xining as programme manager and Kavita as training lead. Both will be supported throughout by our specialists in web design and programme evaluation. We will be drawing on our highly experienced team of training professionals for online support and programme delivery. The final training team will be confirmed on finalisation of the programme agenda with Sakura Bank K.K.

Pricing

The pricing in the table below covers the delivery of all programme components described above. As discussed, these costs are based on using your training facilities in Osaka, Japan. The pricing excludes flights, accommodation and expenses for the Kloud delivery team during residential training but includes all other programme component expenses. Subsequent programme iterations would not incur programme design costs. Payment is in two parts: €16,500 upon contract commencement and €32,500 upon delivery of final report. Cost excludes any applicable taxes.

Programme component	Total cost
Programme design	€10,500
Pre-programme activities	€6,000
Residential training	€30,000
Evaluation and reporting	€2,500
Total	**€49,000**

Figure 15.18 Extract from draft proposal for Sakura Bank K.K. operations improvement executive development

Table 15.5 Chao's notes on the Sakura Bank K.K. project activities and costs

Programme component	*Activity*	*Optimistic estimate*	*Likely estimate*	*Pessimistic estimate*	*Immediate predecessors*	*Daily cost (€)*
Programme design	1. Client kick-off meeting (online)	1	1	1	n/a	500
	2. Training needs analysis	3	3	6	1	500
	3. Trainer contracting	2	3	4	2	150
	4. Programme design	4	5	8	2	500
	5. Client review meeting	1	1	1	3, 4	500
	6. Internal kick-off meeting	1	1	1	5	450
	7. Training material creation	6	8	14	6	350
Pre-programme activities	8. Website set-up	5	5	7	6	250
	9. Website go live	1	1	2	7, 8	250
	10. Pre-programme activities	9	10	14	9	350
Residential training	11. Programme administrative arrangements	8	10	12	6	150
	12. Residential training programme	15	15	15	10, 11	1,500
Evaluation and reporting	13. Post-course survey	3	4	5	12	150
	14. Final report	2	2	4	13	500
	15. Project closure	2	2	3	14	250

typically work on the basis of a 20 per cent mark-up between our costs and the price we charge the client. I think that the margin will be pretty tight on this one.'

Maria left the Managing Director's office and headed for her desk. Tao's final words were ringing in her ears: '*Meet me tomorrow so we can prepare for the kick-off meeting next Monday*.' Sitting down, she looked back over the notes she'd made. Where to begin?

Questions

Based on the information you have, develop a project plan for the Sakura Bank K.K. operations improvement training programme, to share with Tao, the Managing Director of Kloud BV. This should include:

- *Project timing* – complete a critical path analysis, create a Gantt chart and consider any uncertainties in time estimates.
- *Project costing* – create a project budget and consider options for 'crashing' activities.
- *Project resourcing* – create an RACI matrix to determine the key responsibilities for those involved in the project.
- *Project risk* – note any risks you are concerned about and possible mitigation strategies.

Applying the principles

Some of these exercises can be answered by reading the chapter. Others will require some general knowledge of business activity and some might require an element of investigation. **Model answers for the first two questions are available on the student companion website**. *For model answers to the other questions in this section, please ask your tutor/lecturer.*

1. Revisit the Vasa project case example in this chapter.

a) Who should be held responsible for this disaster?

b) What can be learned from the Vasa story for the management of different kinds of modern-day projects?

2. '*Funding comes from a variety of sources to restore the literally irreplaceable buildings we work on. We try to reconcile historical integrity with commercial viability and rely on the support of volunteers. So, we need to involve all stakeholders all the way through the project*.' (Janine Walker, Chief Project Manager, Happy Heritage, a not-for-profit restoration organisation) Her latest project was the restoration of a 200-year-old 'poorhouse' as a visitor attraction, originally built to house local poor. Janine's team drew up a list of stakeholders and set out to win them over with their enthusiasm for the project. They invited local people to attend meetings, explained the vision and took them to look round the site. Also, before work started Janine took all the building staff on the same tour of the site as they had taken other groups and the VIPs who provided the funding. '*Involving the builders in the project sparked a real interest in the project and the archaeological history of the site. Often, they would come across something interesting and tell the foreman, who would involve an archaeologist and so preserve an artefact that might otherwise have been destroyed. They took a real interest in their work, they felt involved*.'

a) Who do you think are the main stakeholders for this project?

b) What are the benefits and risks of involving these stakeholders in the project?

3. The table below shows the activities, their durations and predecessors for designing, writing and installing a bespoke computer database for a commercial bank headquartered in Singapore. Draw a network diagram (activity-on-node) for the project and calculate the fastest time in which the operation might be completed.

Bespoke computer database project activities

Activity	*Duration (weeks)*	*Activities that must be completed before it can start*
1. Contract negotiation	1	-
2. Discussions with main users	2	1
3. Review of current documentation	5	1
4. Review of current systems	6	2
5. Systems analysis (a)	4	3,4
6. Systems analysis (b)	7	5
7. Programming	12	5
8. Testing (prelim)	2	7
9. Existing system review report	1	3,4
10. System proposal report	2	5,9
11. Documentation preparation	19	5,8
12. Implementation	7	7,11
13. System test	3	12
14. Debugging	4	12
15. Manual preparation	5	11

4. Examine this simple domestic project: starting at 6.00 a.m., to make breakfast in bed consisting of a boiled egg, toast and orange juice, using the minimum staff resources and time, and to a high quality (egg freshly boiled, warm toast, etc.). The activities involved in the project, resources and times are shown in the table below. Draw a chart that shows your recommended start and finish times for each activity

Activity	*Resource (person)*	*Duration (minutes)*
Butter toast	1	1
Pour orange juice	1	1
Boil egg	0	4
Slice bread	1	1
Fill pan with water	1	1
Bring water to boil	0	3
Toast bread	0	2
Take loaded tray to bedroom	1	1
Fetch tray, plates, cutlery	1	1

5. The table below shows the planned time and budget for a legal consulting project being developed for a client in Copenhagen, Denmark. Complete an earned value analysis (EVA) for the project based at the end of month 4, given that only activities A, B and C have actually been completed and spending to date has been €38,250.

Activity	*Planned time*	*Planned budget*
A	1 month	€26,000
B	1 month	€10,500
C	1 month	€6,750
D	1 month	€13,000
E	1 month	€9,650
F	1 month	€12,750
G	1 month	€8,750
TOTAL	**7 months**	**€87,400**

6. In the oil industry, project teams are increasingly using virtual reality and visualisation models of offshore structures that allow them to check out not only the original design but any modifications that have to be made during construction.

a) Why do you think a realistic picture of a completed project helps the process of project management?

b) Why are such visualisations becoming more important?

Notes on chapter

1 The information on which this example is based is taken from: Howell, E. and Hickok, K. (2020) 'Apollo 13: The moon-mission that dodged disaster', Space.com, 31 March; Whiting, M. (ed) (2018) 'The legacy of Apollo 6', NASA, 4 April; Taylor Redd, N. (2019) 'Apollo 11: First men on the Moon', Space.com, 9 May; Overbye, D. (2017) 'Cassini flies towards a fiery death on Saturn', *New York Times*, 8 September; Rincon, P. (2017) 'Our Saturn years – Cassini-Huygens' epic journey to the ringed planet, told by the people who made it happen', BBC, 14 September; Cobb, W. (2019) 'How SpaceX lowered cost and reduced barriers to space', The Conversation, 1 March.

2 The information on which this example is based is taken from: Metz, R. (2019) 'The world's biggest spice company is using AI to find new flavors', CNN Business, 5 February; Wiggers, K. (2019) 'IBM and McCormick blend new seasonings with AI', Venture Beat, 4 February; Lougee, R. (2019) 'Using AI to develop new flavor experiences', IBM Research Blog, 5 February.

3 The information on which this example is based is taken from: Schuetze, C. (2020) 'Berlin's newest airport prepares for grand opening. Again', *New York Times*, 29 April; CAPA Centre for Aviation (2020) 'Berlin Brandenburg Airport's terminal certified for opening – at last', centreforaviation.com, 8 May; L.R.S. (2017) 'Why Berlin's new airport keeps missing its opening date', *The Economist*, 25 January.

4 This example was written and kindly supplied by Professor Mattia Bianchi, Stockholm School of Economics.

5 The information on which this example is based is taken from: Burgess, M. (2018) 'Ocado's collaborative robot is getting closer to factory work', *Wired*, 11 January; Butler, S. (2018) 'Ocado to wheel out C3PO-style robot to lend a hand at warehouses', *The Guardian*, 11 January.

Taking it further

There are hundreds of books on project management. They range from the introductory to the very detailed, and from the managerial to the highly mathematical. Here are five general project management books that are worth a look and two journal articles examining more specific aspects of projects:

Cole, R. and Scotcher, E. ***(2015) Brilliant Agile Project Management: A practical guide to using agile, scrum and kanban, Pearson Business.*** *A practical and modern take on project management.*

Davies, A. ***(2017) Projects: A very short introduction, Oxford University Press.*** *A very well-written book on project management, covering key ideas in a very concise way.*

Davies, A. et al. ***(2017) 'Five rules for managing large, complex projects', MIT Sloan Management Review, 31 August.*** *An interesting article focusing on recent research on mega-projects, but giving useful management insights for all large-scale projects.*

Kogon, K., Blakemore, S. and Wood, J. ***(2015) Project Management for the Unofficial Project Manager, BenBella Books.*** *A short, easy-to-read book that helps lay a solid foundation in project management and does a good job bringing in the 'people' aspect of projects.*

Maylor, H. ***(2010) Project Management, 4th edition, Pearson.*** *A very good introductory text on the subject that takes the reader through the key steps of projects.*

Pinto, J.K. ***(2019) Project Management: Achieving competitive advantage, Global edition, Pearson.*** *Long-running text with comprehensive coverage of project management.*

Whyte, J. ***(2019) 'How digital information transforms project delivery models', Project Management Journal, 50 (2), pp. 177–194.*** *An interesting article exploring the implications of digitisation on project management.*

Index

Credits

Text

23 Cambridge University Press: Jensen, M. C. (2001) 'Value maximization, stakeholder theory, and the corporate objective function', *Journal of Applied Corporate Finance*, vol. 14, no. 3, pp. 8–21; **23 Michael C. Jensen:** Michael Jensen, Harvard Business School; **33 Institute for Operations Research and the Management Sciences (INFORMS):** Boudreau, J., Hopp, W., McClain, J.O. and Thomas, L.J. (2003) 'On the interface between operations and human resources management', *Manufacturing & Service Operations Management*, vol. 5, no. 3, pp. 179–202; **45 Sir Martin Sweeting:** Sir Martin Sweeting, PhD, University of Surrey, UK; **51 Communications of the Association for Information Systems :** Osterwalder, A., Pigneur, Y. and Tucci, C. (2005) 'Clarifying business models: origins, present and future of the concept', *CAIS*, vol. 15, pp. 751–775; **52 Business Model Alchemist:** Osterwalder, A. (2005) 'What is a business model?', http://business-model-design.blogspot.com/2005/11/what-is-business-model.html; **53 Cap Gemini:** Based on the definitions developed by Cap Gemini; **58 Kogan Page:** (Case Example) Based on an example from Slack, N. (2017) The Operations Advantage, *Kogan Page*. Used by permission; **63 Academy of Management:** An idea proposed by Jay Barney. See Barney, J.B. (2001) 'Is the resource-based theory a useful perspective for strategic management research? Yes', *Academy of Management Review*; **64 Kogan Page:** (Figure 2.8) From Slack, N. (2017) The Operations Advantage, *Kogan Page*, reproduced by permission; **69 Harvard Business Publishing:** Skinner, W. C. (1969) Manufacturing - Missing Link in Corporate Strategy, *Harvard Business Review*, May 01; **84, 123, 124, 133, 138, 397 Pearson Education:** (Figure 3.2, Figure 4.3, Figure 4.4, Figure 4.6, Figure 4.9 and Figure 11.7) From Slack, N. Operations Management, 8e, © 2016 Pearson Education Limited, UK; **94 SAGE Publications:** Anderson, E., Lin, S., Simester, D. and Tucker, C. (2015) 'Harbingers of failure', *Journal of Marketing Research*: October 2015, vol. 52, no. 5; **106-107 Simon & Schuster:** Based on one presented by Hayes, Wheelwright and Clark in Hayes, R.H., Wheelwright, S.C. and Clark , K.B. (1988) Dynamic Manufacturing, *The Free Press*; **108 Harvard Business Publishing:** Sobek, D.K. II, Liker, J.K. and Ward, A.K. (1998) 'Another look at how Toyota integrates product development', *Harvard Business Review*, July-August; **114 Ellen MacArthur:** Quote by Ellen MacArthur; **120 Pearson Education:** Christopher, M. (2011) 'Logistics and supply chain management', *Financial Times Prentice Hall*, London; **121 Scott Fitzgerald:** Quote by Scott Fitzgerald; **130 Bill Clinton:** Bill Clinton (American 42nd US president (1993-2001)); **132 Harvard Business Publishing:** Moore, J. F. (1993) 'Predators and prey: a new ecology of competition', *Harvard Business Review*, May–June; **137 The Institution of Engineering and Technology:** Bacon, G. Machan, I. and Dnyse, J. (2008) 'Offshore challenges, manufacturing', *Institute of Electrical Engineers*, January; **147 TechTarget:** Definition from techtarget.com; **166 Dishang Group:** Quoted in Dishang Group's website, www.dishan-group.com; **195 Yam Kum Weng:** Mr Yam Kum Weng, Executive Vice President of CAG; **206 Harvard Business Publishing:** Shostack, L. G. (1984) 'Designing services that deliver', *Harvard Business Review*, vol. 62, no. 1, 133–139; **255 Professor Mark Johnson:** Professor Mark Johnson, Warwick Business School; **255 Warwick Business School:** (Figure 7.8) The 50,000 mile journey of Wimbledon's tennis balls, WBS News 02/07/2014, © 2017 Warwick Business School; **262 Michael J. White:** Michael J. White, CEO of the Joint Venture; **262 Martin Gnass:** Martin Gnass, Hapag-Lloyd; **264-266 Jas Kalra/Jens Roehrich/Brian Squire:** (Case Study) This case was co-authored with Jas Kalra, Bartlett School of Construction and Project Management, University College London and Jens Roehrich and Brian Squire, School of Management, University of Bath; **295 BBC:** BBC News site (2017) 'United CEO says removed passenger was "disruptive and belligerent"', 11 April; **298-302 Vaggelis Giannikas:** (Case Study) This case was co-authored with Vaggelis Giannikas at School of Management, University of Bath; **341 NorthWest Airlines:** Farman, J. (1999) 'Les

Coulisses du Vol', Air France. Talk presented by Richard E. Stone, NorthWest Airlines at the IMA Industrial Problems Seminar, 1998; **350 IDG Communications, Inc.:** Kanaracus, C. (2008) 'Waste Management sues SAP over ERP implementation', InfoWorld, 27 March; **386 John Wiley & Sons:** (Table 11.1) Adapted from Fitzsimmons, J.A. (1990) 'Making continual improvement: A competitive strategy for service firms' in Bowen, D.E., Chase, R.B., Cummings, T.G. and Associates (eds), Service Management Effectiveness, Jossey-Bass; **391 Janina Aarts/Mattia Bianchi:** (Case Example) Example written and supplied by Janina Aarts and Mattia Bianchi, Department of Management and Organization, Stockholm School of Economics; **395 Edward Kay:** Quoted by Edward Kay; **419 Macmillan Publishers:** Carroll, L. (1871) Through the Looking Glass, Penguin Classics, 2003; **422 Harvard Business Publishing:** Kaplan, R.S. and Norton, D. (1996) *The Balanced Scorecard*, *Harvard Business School Press*, Boston; **426 Harvard Business Publishing:** Shenkar, O. (2010) Copycats: How smart companies use imitation to gain a strategic edge, *Harvard Business School Press*; **431 Nicholas Brealey Publishing:** Hammer, M. and Champy, J. (1993) Re-engineering the Corporation, *Nicholas Brealey Publishing*; **432 Harvard Business Publishing:** Hammer, M. (1990) 'Re-engineering work: Don't automate, obliterate', *Harvard Business Review, vol. 68, no. 4*; **436 McGraw-Hill Education:** Pande, P.S., Neuman, R.P. and Cavanagh, R. (2013) The Six Sigma Way, *McGraw-Hill*, New York; **472-473 John Wiley & Sons:** Dale, B.G, Bamford, D. and van der Wiele, T. (eds) (2016) Managing Quality, (6th edn), *John Wiley & Sons*; **474 Akihiro Nikkaku:** Quote by Akihiro Nikkaku; **474 EFQM:** The EFQM Website, www.efqm.org; **475 EFQM:** (Figure 13.9) Reproduced with the permission of the EFQM; **475 SAGE Publications:** Hyman, J. and Mason, B. (1995) Managing Employee Involvement and Participation, *Sage*; **535 John F. Kennedy:** May 1961, President John F. Kennedy delivered a speech to the US Congress; **545 Mattia Bianchi:** (Case Example) This example was written and kindly supplied by Professor Mattia Bianchi, *Stockholm School of Economics*; **558-560 Nigel Spinks:** (Case Study) This case was co-authored with Nigel Spinks, Henley Business School, University of Reading.

Photo

5 (bl) Shutterstock: Chubykin Arkady/Shutterstock; **5 (br) Getty Images:** Mint Images RF/Getty Images; **5 (ml) Alamy Stock Photo:** Albert Karimov/Alamy Stock Photo; **5 (mr) Getty Images:** Alistair Berg/DigitalVision/Getty Images; **5 (tl) Getty Images:** Tetra Images/Brand X Pictures/Getty Images; **5 (tr) Getty Images:** FangXiaNuo/E+/Getty Images; **8 Getty Images:** Maskot/Getty Images; **11 Shutterstock:** Vichy Deal/Shutterstock; **37 Shutterstock:** VidEst/Shutterstock; **45 Alamy Stock Photo:** Paul Fleet/Alamy Stock Photo; **69 Shutterstock:** jax10289/Shutterstock; **98 Shutterstock:** Wongtang/Shutterstock; **112 Shutterstock:** LightField Studios/Shutterstock; **121 Shutterstock:** Gabriele Maltinti/Shutterstock; **135 Shutterstock:** AP/Shutterstock; **167 Shutterstock:** BBC Films/Focus Features/Kobal/Shutterstock; **169 Shutterstock:** Shutterstock; **187 Getty Images:** Thomas Barwick/Stone/Getty Images; **195 Getty Images:** John Seaton Callahan/Moment Unreleased/Getty Images; **200 Alamy Stock Photo:** Kristoffer Tripplaar/Alamy Stock Photo; **223 Getty Images:** Tinpixels/E+/Getty Images; **240 Getty Images:** Ruth McDowall/AFP/Getty Images; **245 Getty Images:** Pratan Ounpitipong/Moment/Getty Images; **262 Getty Images:** Jonathan Nackstrand/AFP/Getty Images; **266 Getty Images:** Luke MacGregor/Bloomberg/Getty Images; **278 Alamy Stock Photo:** Chombosan/Alamy Stock Photo; **281 Getty Images:** Metamorworks/iStock/Getty Images Plus/Getty Images; **282 Getty Images:** Naufal MQ/Moment/Getty Images; **300 Shutterstock:** Amritansh Sagar/Shutterstock; **314 Alamy Stock Photo:** American Photo Archive/Alamy Stock Photo; **318 Alamy Stock Photo:** Uli Deck/dpa picture alliance/Alamy Stock Photo; **332 Getty Images:** Helen King/The Image Bank/Getty Images; **335 Alamy Stock Photo:** Charlie Neuman/San Diego Union-Tribune/Zuma Press Inc/Alamy Stock Photo; **341 Shutterstock:** EQRoy/Shutterstock; **354 Shutterstock:** Shutterstock; **356 Alamy Stock Photo:** Imagedoc/Alamy Stock Photo; **388 Alamy Stock Photo:** Cum Okolo/Alamy Stock Photo; **393 Shutterstock:** Disney/Pixar/Kobal/Shutterstock; **410 Getty Images:** Photo_Concepts/Cultura/Getty Images; **424 Getty Images:** Oli Scarff/AFP/Getty Images; **427 Shutterstock:** Shahjehan/Shutterstock; **450 Getty Images:** Klaus Vedfelt/DigitalVision/Getty Images; **462 Getty Images:** JGI/Tom Grill/Getty Images; **466 Shutterstock:** High Level/Shutterstock; **474 Getty Images:** RichLegg/E+/Getty Images; **502 Getty Images:** Ken Cedeno/Corbis Historical/Getty Images; **516 Getty Images:** Michael S. Williamson/The Washington Post/Getty Images; **533 NASA:** JPL-Caltech/NASA; **536 Shutterstock:** monticello/Shutterstock; **543 Shutterstock:** peter jesche/Shutterstock; **551 Shutterstock:** Sorn340 Studio Images/Shutterstock.

Operations and Process Management

Sixth Edition

Master the core subject principles and critically analyse business operations and processes

***Operations and Process Management,* 6th Edition,** by Nigel Slack and Alistair Brandon-Jones, is the market-leading text on operations and process management. This innovative text approaches the subject from a managerial perspective and provides clear and concise coverage of the nature, principles, and practice of operations and process management. The content is organised around diagnostic questions which relate the theory back to the key questions faced by managers in the workplace. Each chapter provides a range of examples, frameworks and techniques designed to help you better analyse existing operations and understand ways to deal with operational and process challenges. This new edition includes new cases studies drawn from a diverse range of industries and regions, and coverage of new topics such as servitisation and the co-creation of services.

This text introduces operations and process management for everyone who wishes to understand the nature, principles and practice of the subject, inspiring a critical and applied mastery of the core principles and processes which are fundamental to successfully managing business operations.

It is aimed primarily at students with some management experience, or those who are about to embark on a career in management.

About the authors

Nigel Slack is an Emeritus Professor of Operations Management and Strategy at the University of Warwick, an Honorary Professor at the University of Bath and an Honorary Fellow of the European Operations Management Association. He is an educator, consultant and writer with wide experience in many sectors.

Alistair Brandon-Jones is a Full Chaired Professor in Operations and Supply Chain Management, and Head of the Information, Decisions, and Operations Division at the University of Bath, School of Management. He is also an Adjunct Professor for Hult International Business School and Danish Technical University.

Pearson, the world's *learning* company.

Front cover image: Jackal Pan/Moment/Getty Images
Cover design by Design Deluxe

Pearson

www.pearson.com/uk

ISBN 978-1-292-35006-6

9 781292 350066 >