

**AN INTRODUCTION TO  
SYSTEMATIC  
REVIEWS**

Sara Miller McCune founded SAGE Publishing in 1965 to support the dissemination of usable knowledge and educate a global community. SAGE publishes more than 1000 journals and over 800 new books each year, spanning a wide range of subject areas. Our growing selection of library products includes archives, data, case studies and video. SAGE remains majority owned by our founder and after her lifetime will become owned by a charitable trust that secures the company's continued independence.

Los Angeles | London | New Delhi | Singapore | Washington DC | Melbourne



DAVID GOUGH  
SANDY OLIVER  
JAMES THOMAS

AN INTRODUCTION TO  
**SYSTEMATIC  
REVIEWS**



Los Angeles | London | New Delhi  
Singapore | Washington DC | Melbourne



Los Angeles | London | New Delhi  
Singapore | Washington DC | Melbourne

SAGE Publications Ltd  
1 Oliver's Yard  
55 City Road  
London EC1Y 1SP

SAGE Publications Inc.  
2455 Teller Road  
Thousand Oaks, California 91320

SAGE Publications India Pvt Ltd  
B 1/1 1 Mohan Cooperative Industrial Area  
Mathura Road  
New Delhi 110 044

SAGE Publications Asia-Pacific Pte Ltd  
3 Church Street  
#10-04 Samsung Hub  
Singapore 049483

---

Editor: Jai Seaman  
Assistant editor: Alysha Owen  
Production editor: Ian Antcliff  
Copyeditor: Christine Bitten  
Proofreader: Derek Markham  
Indexer: David Rudeforth  
Marketing manager: Ben Griffin-Sherwood  
Cover design: Shaun Mercier  
Typeset by: C&M Digitals (P) Ltd, Chennai, India  
Printed in the UK

Editorial Arrangement © David Gough, Sandy Oliver and James Thomas

Foreword © Ann Oakley  
Chapter 1 © David Gough, Sandy Oliver and James Thomas  
Chapter 2 © Rebecca Rees and Sandy Oliver  
Chapter 3 © David Gough and James Thomas  
Chapter 4 © Sandy Oliver, Kelly Dickson, Mukdarut Bangpan and Mark Newman  
Chapter 5 © Ginny Brunton, Claire Stansfield, Jenny Caird and James Thomas  
Chapter 6 © Katy Sutcliffe, Sandy Oliver and Michelle Richardson

Chapter 7 © Jeff Brunton, James Thomas and Sergio Graziosi  
Chapter 8 © James Thomas, Alison O'Mara-Eves, Angela Harden and Mark Newman  
Chapter 9 © James Thomas, Alison O'Mara-Eves, Dylan Kneale and Ian Shemilt  
Chapter 10 © Kristin Liabo, David Gough and Angela Harden  
Chapter 11 © David Gough, Ruth Stewart and Jan Tripney

First published 2017

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act, 1988, this publication may be reproduced, stored or transmitted in any form, or by any means, only with the prior permission in writing of the publishers, or in the case of reprographic reproduction, in accordance with the terms of licences issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to the publishers.

**Library of Congress Control Number: 2016948864**

**British Library Cataloguing in Publication data**

A catalogue record for this book is available from the British Library

ISBN 978-1-4739-2942-5  
ISBN 978-1-4739-2943-2 (pbk)

At SAGE we take sustainability seriously. Most of our products are printed in the UK using FSC papers and boards. When we print overseas we ensure sustainable papers are used as measured by the PREPS grading system. We undertake an annual audit to monitor our sustainability.

# INTRODUCING SYSTEMATIC REVIEWS

DAVID GOUGH, SANDY OLIVER AND JAMES THOMAS



## Aims of chapter

This chapter:

- Introduces the logic and purpose of systematic reviews
- Explains their value for making decisions
- Considers what 'systematic' means when applied to reviewing literature
- Explains how systematic review methods may vary while being systematic
- Introduces some of the current debates
- Explains to readers what to expect from the rest of the book including what is new compared to the 1st edition

## THE ROLE OF RESEARCH REVIEWS

We undertake and read research to find things out. We want to better understand the world. We develop theories and concepts and gather data to develop insights and answer a vast breadth of research questions related to a rich array of disciplines, interests and perspectives of academics, policy makers, professional practitioners, societal groups and individuals. Often we undertake new 'primary' research; but it is also sensible to gather together and examine what is known from existing research. If we do not assess what has been studied already then how can we be in a position to know 'what is known' or to plan what more needs to be studied? This book is about how we can go about this – how to bring together and analyse what we already know from research.

Research needs to use rigorous and accountable methods to be considered research. There needs to be an appropriate systematic rigorous method and this must be explicit so that the results can be interpreted and assessed in the light of how the results were produced. The same logic applies to reviews of research; the reviews need to use appropriate and explicit methods. In other words, reviews of research are a form of research. They are a secondary level of analysis (secondary research) that brings together the findings of primary research. We therefore define a systematic review as 'a review of existing research using explicit, accountable rigorous research methods'. There is not just one way to undertake a systematic review. Research asks all sorts of different questions and uses a wide range of different methods. Reviewing the literature also requires a variety of methods to bring findings together systematically. These range from reviews of statistical data to answer questions about what interventions are effective, to reviews of more qualitative data aiming to develop new theories and concepts.

Reviewing evidence, and synthesising findings, is something that we do all the time when going about our daily lives. For example, consider the range of activities involved in buying a new car. We approach the problem with an overarching question: 'which car shall I buy?' that can be broken down into a series of questions including: 'what cars are available?'; 'what type of car do I need?'; 'which cars can I afford?'; and, if manufacturers' marketing departments are doing their job, 'which car will make me happy?'. We then gather data together to help us make our decision. We buy car magazines, read online reviews, talk to people we know and, when we've narrowed our options down a little, visit car showrooms and take some cars out for a test drive. We critically review the evidence we have gathered (including our personal experience) and identify possible reasons for doubting the veracity of individual claims. If we've decided that we need a small, cheap car, for example, we will understand that the conclusions of a review written by people who like to drive the latest sports cars may be less useful than a review written for the 'thrift supplement' of a weekend newspaper. We may prioritise particular characteristics, such as reliability or boot space, above others, such as fuel economy or safety, and attempt to identify reviews which assess cars with similar requirements to our own.

The example above, while simple compared to the many very complex decisions made in life, introduces us to the purpose of reviews and some of the key issues that we need to grapple with while undertaking a review. Starting our product research by relying first on what other people have written gives us access to a wide range of ideas about how to judge cars, more evidence than we could collect ourselves and leaves us with a smaller task when it comes to visiting showrooms or test driving cars. Our 'decision question' drives what we are doing ('which car shall I buy?') and all the other decisions and judgements we make are based on the need to answer this question. We are faced with many different possible answers (e.g. the make, model and optional extras of our car) and a mass of evidence that purports to answer our question. We need to come to an overall understanding of how this heterogeneous set of data is able to help us come to a decision and, in order to do this, we need to understand why the data are heterogeneous. In the example above, reviews of the same cars come to different conclusions because the people conducting them have different perspectives, priorities and understandings about what they understand the 'best' car to be (and, indeed, what 'best' means). In the same way, reviews often depend on judgements, not only about the methodological quality of research (was it well conducted?), but also its relevance to answering the question at hand.

Our experience as reviewers of research is that there are very many excellent studies published in the social science and health research literatures. However, there are also many studies that have obvious methodological or conceptual limitations or do not report adequate detail for their reliability to be assessed. Even where a study is well conceived, executed and reported, it may by chance have found and reported atypical findings and so should not be relied upon alone. For all these reasons, it is wiser to make decisions based on all the relevant – and reliable – research that has been undertaken rather than an individual study or limited groups of studies. If there are variations in the quality or relevance in this previous research, then the review can take this into account when examining its results and drawing conclusions. If there are variations in research participants, settings or conceptualisations of the phenomena under investigation, these also can be considered and may add strength to the findings (please also see Chapter 8 for a discussion of these issues).

While primary research is essential for producing much crucial original data and insights, its findings may receive little attention when research publications are read by only a few. Reviews can inform us about what is known, how it is known, how this varies across studies, and thus also what is not known from previous research. It can therefore provide a basis for planning and interpreting new primary research. It may not be a sensible use of resources and in some cases it may be unethical to undertake research without being properly informed about previous research; indeed, without a review of previous research the need for new primary research is unknown. When a need for new primary research has been established, having a comprehensive picture of what is already known can help us to understand its meaning and how it might be used.

In the past, individuals may have been able to keep abreast of all the studies on a topic, but this is increasingly difficult and (as we shall see in the next section) expert knowledge of research may produce hidden biases. We therefore need reviews because:

- 1 Any individual research study may be fallible, either by chance, or because of how it was designed and conducted or reported. There are even cases of research reports being fabricated.
- 2 Any individual study may have limited relevance because of its question, scope and context.
- 3 A review provides a more comprehensive and stronger picture based on many studies and settings rather than a single study.
- 4 The task of keeping abreast of all previous and new research is usually too large for an individual.
- 5 Findings from a review provide a context for interpreting the results of a new primary study.
- 6 Undertaking new primary studies without being informed about previous research may result in unnecessary, inappropriate, irrelevant or unethical research.

## SYSTEMATIC, TRADITIONAL AND EXPERT REVIEWS

It has become clear that, when intervening in people's lives, it is possible to do more harm than good (Chalmers 2003) (see Box 1.1). Examining existing research is one way of reducing the chances of doing this. As reviews of such research are increasingly used to inform policy and practice decisions, the reliability of these reviews is critically important. Methodological work on reviewing over the past two decades has built up a formidable empirical basis for reviewing health care evaluations, and systematic reviews have therefore become established as a key component in evidence-informed decision making. So influential has the use of

research through systematic reviews become that their development can be considered to be one of the turning points in the history of science:

This careful analysis of information has revealed huge gaps in our knowledge. It has exposed that so-called 'best practices' were sometimes murderously flawed; and by doing nothing more than sifting methodically through pre-existing data it has saved more lives than you could possibly imagine.<sup>1</sup>

The logic of reviewing is thus twofold: first, that as the opening section discussed, looking at already existing research evidence is a useful thing to do; and second, that as reviews inform decisions that affect people's lives, it is important that they be done well. This book uses the term 'systematic review' to indicate that reviews of research are themselves pieces of research and so need to be undertaken according to some sort of method.

BOX

1 ■ 1

### Examples of decisions not informed by research

**Expert advice:** Dr Benjamin Spock's advice to parents was to place infants on their fronts to sleep – advice not supported by research. When this policy was reversed, rates of sudden infant death dropped dramatically (Chalmers 2001).

**Expert panel:** In the BSE (mad cow) crisis in the UK in the late twentieth century where there were many deaths from eating infected meat, '... highly problematic policy decisions were often misrepresented as based on, and only on, sound science' (van Zwanenberg and Millstone 2005).

**Well-intentioned interventions:** In the Scared Straight programme criminals give lectures to 'at risk' youth about the dangers of a life of crime, but this is statistically associated with higher not lower rates of crime in the at-risk youth (Petrosino et al. 2002, 2013).

Reviewing research systematically involves four key activities: clarifying the question being asked; identifying and describing the relevant research ('mapping' the research); critically appraising research reports in a systematic manner, bringing together the findings into a coherent statement, known as synthesis; and establishing what evidence claims can be made from the research (see Box 1.2 for definitions). As with all pieces of research, there is an expectation that the methods will be explained and justified, which is how we reach our definition that a systematic review is a review of existing research using explicit, accountable rigorous research methods.

Most literature reviews that were carried out a decade or more ago were contributions to academic debates, think pieces, not undertaken in a systematic way. Reviewers did not necessarily attempt to identify all the relevant research, check that it was reliable or write up their results in an accountable manner.

<sup>1</sup>BBC Radio 4, *Moments of Genius*: Ben Goldacre on systematic reviews. Available at: [www.bbc.co.uk/radio4/features/moments-of-genius/ben-goldacre/index.shtml](http://www.bbc.co.uk/radio4/features/moments-of-genius/ben-goldacre/index.shtml).

Traditional literature reviews typically present research findings relating to a topic of interest. They summarise what is known on a topic. They tend to provide details on the studies that they consider without explaining the criteria used to identify and include those studies or why certain studies are described and discussed while others are not. Potentially relevant studies may not have been included because the review author was unaware of them or, being aware of them, decided for reasons unspecified not to include them. If the process of identifying and including studies is not explicit, it is not possible to assess the appropriateness of such decisions or whether they were applied in a consistent and rigorous manner. It is thus also not possible to interpret the meaning of the review findings.

**BOX**

1-2

### Key terms

**Systematic:** undertaken according to a fixed plan or system or method

**Review:** a critical appraisal and analysis

**Explicit:** a clear, understandable statement of all the relevant details

**Accountable:** answerable, responsible and justified

**Map (systematic):** a systematic description and analysis of the research field defined by a review question

**Synthesis:** creating something new from separate elements

**Systematic review:** a review of existing research using explicit, accountable rigorous research methods

**Evidence claim:** the statements that can be justified in respect of answering the review question(s) from the research evidence reviewed

The aim of reviewing systematically is to have such explicit, rigorous and accountable methods. Just as primary research is expected to report transparent, rigorous methods, the same standards can apply to systematic reviews. Just as primary research is undertaken to answer specific questions, reviews of existing research can be productively focused on answering questions rather than addressing topic areas. The focus on a question drives the choice of the methods to find the answers.

Individual experts or expert panels are often consulted to answer questions about what is known from research. Experts may of course have many specialist skills, including knowledge of research, practical experience of the phenomena being considered, and human insight and implicit knowledge that have not been formalised in research. However, there can also be dangers from this richness of knowledge not being explicit.

One danger is that the experts' ideological and theoretical perspectives, and thus the conceptual framework determining their assessment of the research, will not be explicit; and as

with everyone, these perspectives may be influenced by personal interests in the issues being discussed. Second, the boundaries of the experts' knowledge may not be transparent; that is the boundaries of studies familiar to them and thus the evidence being considered. A third danger is that even if the boundaries of the studies are clear, the expert may know some of the studies within those boundaries better than others, so not all the research will have equal representation in the conclusions they draw. Fourth and fifth dangers are the related problems of how the experts assess the quality and appraise the relevance and then synthesise different pieces of evidence. Sixth, it may not be clear the extent to which the expert draws on other forms of knowledge, such as practice knowledge, in forming their overall conclusions. An expert witness in a court may, for example, provide an opinion that the court believes is based on research, but is in fact based on a mixture of research and practice wisdom. Seventh, the manner in which someone is assessed as being expert on a particular area may not be appropriate. They may not be expert at all on this topic, or they may be expert but their esteem and credibility is based on practice knowledge. A court may, for example, give high credibility to research reports from someone who has high esteem as a practitioner rather than as a researcher.

In many ways an expert review or expert panel is similar to a traditional literature review. There may be great insight and knowledge but with a lack of transparency about what this is or how it is being used. With experts and expert panels there may be a lack of clarity about the:

- perspective and conceptual framework, including ideological and theoretical assumptions and personal interest;
- inclusion criteria for evidence;
- nature of the search for evidence;
- sort of evidence that is thus being considered;
- quality and relevance appraisal of that evidence;
- method of synthesis of evidence;
- use of evidence other than research;
- basis for their expertise: (i) how their expertise is assessed, (ii) how its relevance to the topic in question is assessed, (iii) how its relationship to research skills and knowledge is assessed;
- basis of the evidence claims made.

In order to address such issues, systematic reviews often proceed through a number of stages, as shown in Figure 1.1. We shall see in later chapters that these stages are an oversimplification, but they are sufficiently common for them to provide a guide to the part of the review process being discussed in each chapter (the structure of the book and how the chapters fit into this is discussed at the end of this chapter).

Though the idea of systematic reviews is simple, and their impact profound, they are often difficult to carry out in practice, and precisely how they should be done is the subject of much debate and some empirical research. Identifying the relevant research, checking that it is reliable and understanding how a set of research studies can help us in addressing a policy or practice concern is not a straightforward activity, and there are many ways of going about reviewing the literature systematically. This book takes a careful look at how approaches differ (while still being systematic) and considers the issues involved in choosing between these different approaches.

## QUESTIONS, METHODS AND ANSWERS

Primary research asks many different questions from a variety of standpoints (Gough et al. 2009), and this richness of questions, approaches and research methods is also reflected in systematic reviews. If, for example, questions are asked about the meaning attached to different situations by different actors in society, then qualitative exploratory methods are likely to be appropriate. To find out how societal attitudes vary across the country, then a survey may be most helpful, whereas knowing how many people receive different state services may be found from routine government administrative data. To investigate whether a particular social intervention has had the required effect, an experimental study may be the most powerful method.

The idea that different research questions may be answered best by different methods and by different types of data also applies to reviews. For instance, systematic reviews addressing questions about the effects of health interventions have widely agreed systematic methods for setting the scope, judging the quality of studies, and presenting the synthesised findings, often using statistical meta-analysis of the results of randomised controlled trials (see, for example, Higgins and Green 2011). However, a systematic question-driven approach to reviews can apply equally to research questions of process or of meaning that are addressed by more qualitative primary research and by review methods that reflect those qualitative research approaches (see, for example, Patterson et al. 2001; Pope et al. 2007; Sandelowski and Barroso 2007).

Reviews and their findings can vary on many different ‘dimensions of difference’ and these are discussed in detail in Chapter 3. The diversity of methods that are used to bring together (‘synthesise’) study findings lie on a continuum between approaches that aim to *aggregate* or ‘add up’ findings from multiple, similar studies; and those that aim to *configure* or ‘organise’ findings of studies (Sandelowski et al. 2006, 2011; Voils et al. 2008). Aggregative methods of analysis are often used to answer tightly specified questions using quantitative pre-specified methods to test theory using empirical observations. Configurative approaches are more likely to address more open questions that explore and explain variation in study findings. While some configurative methods of analysis bring together qualitative data with an emphasis on the exploration of meaning and people’s lived experiences, many statistical methods also configure study findings to understand why a given ‘dependent variable’ might vary in different situations (see Chapters 8 and 9). Most reviews involve some aggregation and some configuration.

Reviews need to specify the questions they are asking and the methods used to address these, and this is often written as a ‘protocol’ prior to undertaking the review (see also Chapter 4). Writing a protocol or other specification of the review methods at the beginning of a review can be a very useful activity, as it helps the review team to gain a shared understanding of the scope of the review and the methods that they will use to answer the review’s questions. It can also form the basis of the review report. Other benefits thought to accrue from writing a protocol include a reduction of bias (*PLoS Medicine* Editors 2011), though this area is somewhat contested, as others would say that protocols are inappropriate for some types of review where there is a lot of iteration in the method during the process of the review and iterative methods are required at some stages – for example, searching for relevant studies (Greenhalgh et al. 2005a, 2005b).

BOX

1-3

### Key terms

**Quantitative research:** the systematic empirical investigation of quantitative properties of phenomena and their relationships. Quantitative research often involves measurement of some kind.

**Qualitative research:** in-depth enquiry to understand the meaning of phenomena and their relationships.

**Aggregative synthesis:** where the synthesis is predominantly aggregating (adding up) data to answer the review question. Aggregation commonly uses quantitative data but qualitative data can also be aggregated.

**Configurative synthesis:** where the synthesis is predominantly configuring (organising) data from the included studies to answer the review question. Qualitative and quantitative data can be configured to understand variation in study findings. Aggregation and configuration fall on a continuum, and all reviews are likely both to aggregate and configure data to some extent.

**Protocol:** a statement of the approach and methods to be used in a review made prior to the review being undertaken.

With these complex choices about review methods, it is important for reviewers to have a clear understanding of the ‘meaning’ of the question. For example, the starting point for a review might be, ‘We want to know about the effects of classroom teaching assistants’. Further discussion may clarify that this question could be more precisely framed to indicate the more specific question being asked and then the type of synthesis that would be most appropriate, for example:

- ‘Do students in classes where there is a classroom teaching assistant get higher or lower scores on test scores?’ The most appropriate synthesis method is probably aggregative quantitative, preferably meta-analysis (if the primary research meets the necessary conditions).
- ‘How can we conceptualise the way that the presence of classroom assistants changes relationships between students and teachers and between teachers in class?’ The most appropriate synthesis method is probably configurative (if the primary research meets the necessary conditions).

Once the meaning of the question is clear, the appropriate method can be chosen. Although it may seem a simple task to match methods to question types, the reality is more complex. A question may be answered in more than one way. The selection of the appropriate method for a review depends on the type of question to be answered, but also must consider the use to which the review will be put, and practical issues, such as the experience and perspective of the review team (see Chapter 4). Also, reviews vary in how extensive the review question is (the breadth of question and the depth in which it is examined) and in the time and resources used to undertake it (see Chapter 3).

When thinking about the role of a review, it can be helpful to always keep in mind the basic logic of undertaking reviews:

- What is the issue at hand?
- What do we want to know (from research that might help us with this issue – and so what are our research questions)?
- What do we know already (so undertake or use existing systematic reviews)?
- And how do we know it (from what methods of systematic review and what type of primary research)?
- What more do we want to know (from further reviews or further primary research)?
- And how could we know it (how can the systematic reviews inform further research strategies)?

## CHALLENGES OF, AND CRITIQUES OF, SYSTEMATIC REVIEWING

Systematic reviewing has only recently become a major area of methodological development. Although the idea of being more explicit about reviewing research is not new, it was only in the 1980s that some of the texts on some of the major types of review, such as statistical meta-analysis (Glass et al. 1981) and meta-ethnography (Noblit and Hare 1988), were published (these types of review are explained later in the book, particularly in Chapters 3 and 8). Although reviews of literature have been advocated for very many years (Bohlin 2012; Chalmers et al. 2002; Shadish and Lecy 2015), systematic reviewing is still a young and rapidly developing field of study and methods of reviewing have not yet been developed for all areas of science. It is an exciting time yet there are many challenges to be overcome (Oakley et al. 2005).

First, there are many conceptual and methodological challenges. This book discusses a wide range of approaches to reviewing from an equally broad range of different epistemological positions. In it we argue that this range of methods is useful, but we realise that this diversity raises many complex issues, particularly in relation to mixing results from different research traditions. Related to the conceptual challenges are more detailed methodological issues. While there is a strong empirical base underpinning some review methods, many have been designed and developed based on primary research methods and on the logic of systematic reviews, with comparatively little methodological study of the impact of different approaches to particular aspects of reviewing – such as search strategies and data coding. The methods of evidence-informed policy and practice are not always evidence informed; we need more empirical data to support the selection of specific review methods.

A second challenge, and related to the methodological issue, is the lack of an agreed terminology to describe, discuss and develop methods (Gough et al. 2012). Some of this linguistic confusion arises from fundamental debates about the nature of knowledge and the role of research in this; and some is a lack of clarity about widely used but unclear distinctions, such as quantitative and qualitative research. There are, however, many further problems in the terminology used for reviews of research evidence.

The term ‘meta’ is one such confusing word. Meta has many meanings and in relation to research it is often used to mean ‘about’ or ‘beyond’ and so ‘meta-evaluations’ are ‘evaluations of evaluations’. Such meta-evaluations can be a form of systematic review, but they can also be simply the evaluation of the quality of one or more evaluations (Gough et al. 2014). The term ‘meta-analysis’ can mean analysis of analysis and so can be

another term for systematic review, but meta-analysis has been used so often to refer to statistical reviews that in some circles it has become synonymous with statistical synthesis. A 'meta-review', on the other hand, is a review of ('about') reviews, of which there are several forms (see Chapter 3). Also, many words used in reviewing can give the impression that a particular type of review is being assumed. Words such as 'protocol' (methods of a review) suggest pre-specified methods and the word 'synthesis' may suggest aggregation to some people and the configuring of findings to others (see Chapter 3 on 'dimensions of difference' in reviews). This book does not provide an overarching classification for systematic reviews. Instead it aims to discuss the nature of reviews and how they vary and includes some of the terminology used to describe this variation (see Chapter 3).

**BOX**

1-4

**Key terms**

**Meta:** about, above or beyond something.

**Meta-evaluation:** evaluation of evaluations. These can be systematic reviews; alternatively, they can be formative or summative evaluations of evaluations, including standards for such evaluations.

**Meta-analysis:** the term has been used in the past to refer to all forms of review, but now usually refers to the statistical meta-analysis of data from primary research studies.

**Review of reviews:** a review of previous reviews. This contrasts with reviews that only review primary research studies.

A third challenge relates to resource constraints. Reviews are major pieces of research and require time and other resources. The resources available will impact on the type of review that can be undertaken. More fundamentally, there is the extent of investment by society in reviews of research. This is not just an issue of overall funding for research, but the balance of investment between: primary research; reviews of what is known from that primary research; and systems to assist the interpretation and application of that research (see Chapter 11). Currently, the funding of research is predominantly on primary production compared to synthesis and use of research.

The most appropriate balance between primary research and reviews is difficult to specify but is likely to vary for different funders, producers and users of research. A research institute or evidence or brokerage centre closely related to policy or practice decision makers, for example, would be particularly likely to have a high balance of reviews to primary research. The challenge for all the individuals and organisations involved in research is to consider whether their needs, roles and responsibilities are best met by the current balance they take between primary research and reviews of and use of that research.

A fourth challenge is the capacity constraints in terms of individual and organisational skills and infrastructure to undertake reviews. There are relatively few people with advanced review skills and so even if funding was available, it would take time to build up the necessary capacity.

Fifth are the capacity constraints for using reviews. This not only involves the capacity to read and understand reviews, but also the capacity to interpret and apply reviews in meaningful and useful ways. Reviews of research are only one part of the research generation and use cycle. The cycle cannot work effectively without active engagement between research and the users of research and this may require further intermediary processes and intermediary organisations (Gough et al. 2011; and see Chapter 11). If formal processes are required to be explicit about ideological and theoretical perspectives and methods in the production of knowledge and its synthesis, then maybe similar processes are also required to support the use of such knowledge.

Sixth are broader political challenges. There are many critics of systematic reviews. One criticism is the mistaken belief that systematic reviews are only concerned with questions of studies of effectiveness and so represent an empiricist (or positivist) research paradigm. In social science there are often strong views about the appropriateness of different research paradigms (Oakley 2000a, 2000b) and some argue that the empiricist paradigm is deficient, making systematic reviews deficient too. However, as has already been explained in this chapter, the logic of reviews can apply to many questions and methods, not only empirical statistical reviews; meta-ethnography, for example, was introduced as a method in the late 1980s (Noblit and Hare 1988). Moreover, we argue that systematic reviews of effectiveness, framed with the help of stakeholders, are not deficient but important contributions to accumulating knowledge.

A related criticism is that the review process is atheoretical and mechanical and ignores meaning. This is another criticism of the empiricist paradigm where (in both primary research and reviews) a pre-specified empiricist strategy is used to test hypotheses such as the effectiveness of interventions. It is a criticism of a particular research paradigm and of the narrowness of some studies within that paradigm rather than of systematic reviews. The preference of one research paradigm over another or the existence of some poor quality primary research studies or systematic reviews is not an argument about the inherent appropriateness or importance of systematic reviews. Decision makers at various levels need to have different kinds of research questions addressed in order to inform the formulation of policy or practice and to implement change, and so the authors of this book value a plurality of perspectives, research questions and methods and thus also of review questions and methods (Oakley 2000a, 2000b). We value theory-testing reviews asking questions of effectiveness that aggregate findings and we also value reviews that configure and develop and critique concepts and ideas.

Another criticism is that reviews often only consider relatively few studies and thus are ignoring much relevant research. There are at least two issues here. First, many reviews have narrow review questions and so narrowly define the boundaries (inclusion criteria) of the studies they consider and so their conclusions must be limited to research so defined. This needs to be explicit in the title, introduction and summary of a review to avoid misrepresenting the data on which it was reaching its conclusions. The review needs to state what it is not including and thus what it is not studying. A review that was titled 'the importance of X for Y' which only considered a few aspects of X and Y could rightly be criticised for misrepresentation or bias.

The criteria for the inclusion and exclusion of studies can include such issues as the topic focus, the method of primary research and the quality of the research. Researchers with different perspectives and working within different research paradigms will have different views of

what constitutes good quality and relevant evidence. This is the nature of academic discourse and occurs just as much in primary research as with research reviews. With reviews the argument is being played out at a meta-level rather than in discussing individual studies. Reviews allow broader discussions with explicit assumptions and leveraging many studies rather than debates about individual studies.

The second issue in relation to the low numbers of studies in some reviews is concerned with the inefficient process of searching for studies on electronic databases. Many irrelevant studies have to be sifted through to find the few that are on topic. Reports of reviews will include the number of studies found in the search, which may be very many thousands, of which only a few may be relevant. Critics use those numbers of discarded studies to argue that studies are being ignored. What is being ignored here, however, is that electronic searching is imprecise and captures many studies that employ the same terms without sharing the same focus (see Chapters 5 and 7). These extraneous studies need to be excluded after the electronic searching is completed. In sum, there are two related issues: what is the focus of a review and what number of studies will help in addressing that focus?

A broader and potentially more powerful criticism is that systematic reviews appeal to government because they fit with a new managerialism for controlling research. The state can specify what research it wants and how this should be reviewed and thus control the research agenda and the research results. The overt process of setting review questions in discussion with a range of different users is one way to engage with such concern though there are of course issues of who has the power and control to decide perspectives and priorities. Researchers and research funders are in a very strong position to determine the nature and outcome of research. Involving a broader range of users in defining reviews of what we know, what we don't know and what more we want to know, can give voice to others in society and make research more, not less, democratic (Gough and Elbourne 2002; Gough 2007a, 2011; and see Chapter 2). Being more explicit about the personal and political in research and increasing the potential for the increased involvement of different sections of society nationally and internationally is an important goal for all research including systematic reviews.

Systematic reviews have an integral role in the production of research knowledge and are an essential part of the process of interpreting and applying research findings to benefit society. Systematic reviews play a key part in developing future primary research and in advancing methods that better achieve their purpose – so-called 'fit for purpose' research methods. They provide a potential means by which all voices in society can engage in determining research agendas. Whether these voices are heard depends on socio-political forces within and beyond research.

## THE AIMS OF THIS BOOK

This book provides an introduction to the logic of systematic reviews, to the range of current and developing methods for reviewing, and to the consequences of reviewing systematically for the production and use of research. There are many excellent books available on different types of systematic review. This book differs from most others currently available in examining the nature of the basic components of reviews driven by any research questions and including any research methods and types of data.

It examines formal, explicit and rigorous methods for undertaking reviews of research knowledge. This idea of gathering research literature together is straightforward; the challenge is that research questions and methods are very diverse and so we need many different types of review addressing different questions using different types of research for different purposes. Understanding the complexities of reviewing research has practical relevance for using the messages that research offers, but its importance is greater than this.

The book has been designed as a resource for four main audiences. First, it is a resource for those undertaking reviews. It does not provide a step-by-step guide for carrying out every stage of every possible type of review. Considering detailed methods of reviewing for all possible types of research question, ranging from measuring the effects of interventions to developing theory to understand how things can be best understood, would be too much for a single volume and sections would quickly become out of date as new insights are being published regularly. Instead, it aims to provide an explanation of the main issues encountered at different stages of different types of systematic reviewing, and thus the thinking behind the many decisions required in any review. As we shall see later in the book, reviews vary considerably in their questions and methods, and in their scope and purpose. An understanding of how the aims of a review can be achieved is the most fundamental requirement for being able to undertake a review or for using review findings appropriately. The book is described as an introduction because it makes known to readers, in considerable depth, the thinking underlying the aims and methods of reviews.

Systematic reviews raise issues about how primary research is undertaken, how different approaches and methods are fit for purpose, and the implications for what more needs to be known and how the gaps can be filled by primary research. The second audience for the book therefore consists of those who fund, plan or undertake primary research. Reviews tell us what we know and don't know in relation to a question, and how we know this. They also raise the issue of what future research might be, what we might know and how we might know it. The review process thus enables a consideration of what would be the appropriate, fit-for-purpose research strategies and methods to achieve specific research objectives. It also provides an opportunity for non-researchers to be involved in such processes; to consider the research to date and to participate in developing future research agendas.

The third audience for the book consists of those who may use reviews to inform decision making. If decisions are being made on the basis of a review, and yet reviews can vary considerably in terms of focus, question, purpose, method, data and rigour, then an understanding of these characteristics is important for interpreting the quality and relevance of that review for the decision makers.

The fourth audience for the book will be those who have a wider interest in the production and use of research in society. In being question-driven, systematic reviews raise issues about the purpose of research and thus the drivers producing that research. They raise questions about whose questions are being asked and the methods being used to provide those answers. Reviews of research are driven by the needs of the different people asking these different questions, including all those who may use the research or other stakeholders affected by it. For this reason, reviews and their methods need to be relatively fit for achieving their intended purpose(s). These questions relate to fundamental issues about the funding and use of research and its role in society. They also relate to how reviews can be interpreted and applied in practice, thereby raising a whole range of issues about moving from knowledge to action, described in different ways as knowledge translation, knowledge mobilisation and exchange (see Chapter 11).

The book reflects the development of theory and practice at the EPPI-Centre where all of the authors do or, until recently, did work (Oakley et al. 2005). There is not an EPPI-Centre method for undertaking reviews; rather there are principles that guide our work. These include the following:

- 1 Both primary research and reviews of research are essential to advance knowledge.
- 2 There is a wide range of review methods just as there is a wide range of primary research methods.
- 3 Reviews should follow the research method principles of quality, rigour and accountability that are expected in primary research.
- 4 Review methods often reflect the methods, epistemological assumptions and methodological challenges found in primary research.
- 5 Reviews should be driven by questions which may vary in many ways, including their implicit or explicit assumptions.
- 6 Those asking questions of research are ‘users’ of that research, with particular perspectives (values with associated priorities and ideological and theoretical assumptions) that inform the process of undertaking primary and secondary (reviews of) research.

Although the main focus of the work of the EPPI-Centre has been on the development of methods of systematic review, we are increasingly working on the very related area of the use of reviews in decision making and ‘research on research use’ (see Chapter 11). This is also reflected in our teaching which includes both a MSc in Systematic Reviews for Policy and Practice and a MSc in Social Policy and Social Research.

## DEVELOPMENTS IN THIS 2ND EDITION

This 2nd edition of the book has the same overall aims and approach as the 1st edition (Gough et al. 2012a), but as five years have passed since the 1st edition every chapter has been updated and revised.

The main changes in structure are:

- The expansion of the text on synthesis from one to two chapters (Chapters 8 and 9). There is simply too much material to include in a single chapter. Chapter 8 discusses the broad range of approaches to synthesis in terms of the aims of synthesis and how this is achieved; and focuses on non-statistical methods of synthesis. Chapter 9 provides more detail on technical aspects of some quantitative approaches to synthesis.
- The chapter on critical appraisal of studies included in a review has been changed to be one aspect of a larger chapter considering the evidence claims made by a review. This is partly because critical appraisal occurs throughout a review, not just a specific stage. More fundamentally, it is because critical appraisal is also undertaken on the whole review in terms of the quality and relevance of its methods, the sufficiency of its data from included studies, and the extent that it is able to make evidence claims to answer the review question. The chapter has thus been re-named as ‘Developing Justifiable Evidence Claims’ and moved to Chapter 10. In other words, moved from being before to being after the chapters on synthesis (Chapters 8 and 9).
- The final chapter of the 1st edition, on future developments has been removed. Instead, individual chapters discuss how methods are developing.

There are too many specific changes of detail to list here but there are a few aspects of the 1st edition that we have been keen to emphasise more in this 2nd edition.

First, we continue to stress that systematic reviews are not limited to any particular research question or method and that reviews will vary in aim and method just as with primary research. Systematic reviews differ in some ways from primary research but for us the best way to consider them is as a level of analysis rather than being seen as a very different form of undertaking to primary research.

Second, we continue to emphasise the way that the rich diversity of reviews varies on a number of key dimensions (Gough et al. 2012b). There is an increasing growth of brands of types of reviews which we welcome in terms of the growing rich diversity of reviews. However, without specifying where these review types sit on the key dimensions there can be a lack of clarity about the distinctive characteristics of these 'brands'; and even when a review might be given a particular brand 'label', there is such wide variation in the range of approaches covered by a label, that it does not necessarily describe the approach taken very precisely.

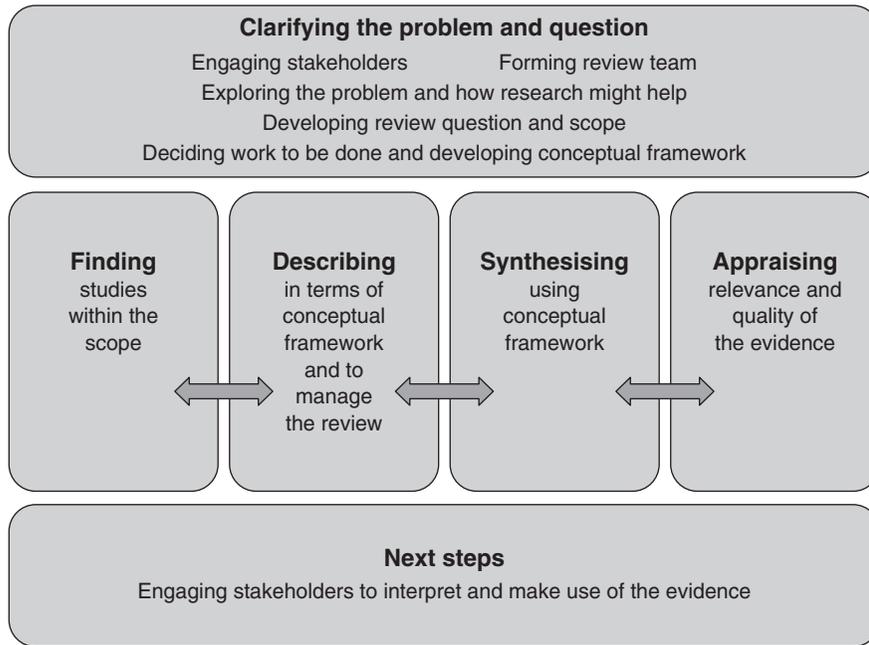
Third, we have made some adjustments to the 'dimensions of difference' we use to describe variation between reviews (Gough et al. 2012a, 2012b). We have adjusted the language of the terms we use and introduced new dimensions such as the extent to which methods adopt widely agreed procedures (see Chapter 3). Also, in the 1st edition, we highlighted the distinction between aggregative and configuring methods of analysis in reviews. In this second edition, we have increased our discussion of configurative approaches and the way in which configuration and aggregation are often both present in different ways in many types of research synthesis.

Fourth, we have put even more emphasis on considering the extent of the 'research problem' that a review attempts to address. Both primary research and reviews of that research vary in their ambition and resources and how much they aim to achieve. We therefore frame reviews as a process that takes place between: (i) there being an issue that somebody wants to know more about from research; and (ii) what the review achieves in answering that question. We consider reviews in relation to the user perspectives, aims of a review, and the role that reviews play in the research production and use process. This approach is shown in the diagrams at the start of each chapter showing its position as prior to, after, or stages within a review.

Fifth, there are some general changes in research that have parallel developments in reviews. One example is the advances in information technology that is changing the relationship between individual level, primary research level and synthesis level of data analysis. Another related example is the increased use of large data sets including routine administrative data and the blurring between experimental and observational research. Another example is the increased awareness of the need to consider theory and empirical data together with more complex primary and consequently more complex secondary forms of research analysis (see Chapters 8 and 9). In addition, developments in information technology have also supported the advance of tools and technologies of information management in reviews (see Chapter 7).

## STRUCTURE OF THE BOOK

The book is organised around the stages prior to, during or after a review as illustrated in Figure 1.1 which is used at the very start of most of the chapters of the book (indicating which stage is being considered in each chapter).



**Figure 1.1** Stages of the review process

The current chapter (Chapter 1) provides an introduction to the book. Chapters 2, 3 and 4 are concerned with the initial stages of clarifying the problem and question, and how this will be addressed in the review.

Chapter 2 (Stakeholder Perspectives and Participation in Reviews) focuses on the multiple perspectives that can be involved in a review reflecting different priorities, values and theoretical assumptions. It discusses how these effect a review and how such different perspectives can be brought into and participate in the review process. Chapter 3 (Commonality and Diversity in Reviews) considers the range of different types of review and how they vary. It examines the nature of this variation on such 'dimensions of difference' as structure, resources and approach of the review. Chapter 4 (Getting Started with a Review) then examines how to begin the process of a systematic review. It discusses in more detail the choice of aim and method of review, the team to undertake the review, quality assurance of the review process and product, and the management of all this work.

Chapters 5 and 6 follow on with the next stages of a review. Chapter 5 (Finding Relevant Studies) explores the work required to find the research that will be included in a review. This includes the overall approach and specific strategies and resources and 'screening' to check that the studies do meet the needs of the review. Chapter 6 (Describing and Analysing Studies) follows by explaining the different aims and methods of identifying and recording (coding) relevant information from the studies included in a review. This includes the principles and processes of good coding and how these vary between different types of review.

A review requires the management and analysis of many pieces of detailed information. Chapter 7 (Tools and Technologies for Information Management) examines the ways that information technology can assist the review process in terms of supporting individual stages

such as data coding and analysis as well as the management of the whole review, and the relationship between different levels of analysis in research (from raw data, to analysis in primary research and reviews of such research).

Chapters 8 and 9 then examine the methods used to synthesise the findings of included studies. Chapter 8 (Synthesis Methods for Combining and Configuring Textual or Mixed Methods Data) develops the ideas introduced earlier in Chapters 3 and 4 on the variation between reviews to consider how reviews can vary in their relation to theory and conceptual development. The chapter also introduces a range of methods of non-statistical and mixed methods synthesis. Chapter 9 (Synthesis Methods for Combining and Configuring Quantitative Data) then concentrates on quantitative methods of analysis such as statistical meta-analysis.

The final stage of the review process is considered in Chapter 10 (Developing Justifiable Evidence Claims). The chapter considers the three main components for assessing the extent that a review has addressed the question that it aimed to answer (the methods of the review, the studies included in the review and the totality of evidence produced by that method using that included evidence).

The final chapter of the book, Chapter 11 (Using Research Findings) focuses on how the findings of a review might be used. It provides a discussion of the nature of research use, how this occurs in practice, how this use might be developed and changed and how we can study these processes of research use. The chapter is relevant to the whole book as it provides some background as to how reviews might be used and so inform the development of methods to assist such use of research.

## CONCLUSION

The starting point for the book is that research as systematic enquiry is an important form of knowledge and that we should balance the investment of resources and energy in new research with the reviewing of what we know from previous research and processes for interpreting and applying that knowledge. Reviews should be one of the first things that we do before embarking on any form of important decision including academics planning new primary research. Traditionally, reviews have been undertaken without clear, formal, explicit and systematic methods, which undermines their status and utility as research, and similar arguments can be made about expert opinion (even if such opinion may have great uses in other circumstances). Systematic reviews have common principles and similar processes, but can vary as much as primary research in terms of their extent, breadth and depth, and in types of question, data and method. Systematic reviews, like any form of research, can be undertaken well and badly, and so need appropriate quality assurance processes to evaluate them. To progress, systematic reviewers need to be aware of the many practical, methodological and political challenges involved in this work and their wider role in the production and use of research in society.