1 Different types of evidence review

Every year a vast number of health-related research studies are carried out. There are randomised controlled trials of new treatments, surveys of patient experiences, evaluations of interventions designed to improve health and/or reduce health inequalities, analyses of routinely collected hospital episode statistics, economic evaluations of the costs and effectiveness of health care, case studies of new practice, research on the experience of service users, patients and health care professionals and so on. This research is conducted by researchers from a diverse range of disciplinary and professional backgrounds, including statistics, epidemiology, psychology, sociology, anthropology, political and economic science, geography, public health and the medical, nursing and therapy professions, and it draws on a variety of different theoretical and methodological approaches. As a result of all this research activity, each week journal papers, books and reports add to the already vast number of publications about the effects and impacts of health interventions, the experience of illness and health care, and the organisation and reorganisation of health-related services. These studies are published in social science, medical, nursing and health-related journals, in books and monographs, and increasingly as electronic publications, online. Alongside this published literature there is also a large but less visible ‘grey’ or unpublished literature which includes research reports, policy briefings and unpublished dissertations. There are also other types of evidence which might inform policy- and decision-making such as consultations with expert and lay groups, and a host of different types of information available via the World Wide Web.

There is growing awareness of the need to strengthen the link between knowledge derived from this mountain of diverse evidence and the decisions made in health care policy and practice (Haines et al. 2004). Surely all this evidence can tell us what should be prioritised, funded, and developed? Unfortunately there are problems with relying on this evidence. Some studies may be too small to provide reliable and valid evidence. Other evidence may be context-bound – specific to a particular country, a region or much smaller
unit of analysis such as a group of patients – such that it is difficult to see which findings are most salient and/or how the findings might be applied in another context. Some research is biased by flaws in the study design or methods, and this makes using the findings highly problematic. Research may tackle only one small part of the question or problem which decision-makers want to know about: unpublished studies and other sources of evidence need to be drawn on. There are often additional problems with determining the quality and rigour of non-research sources of evidence.

This information mountain, and the methodological and other problems associated with it, presents major challenges for policy-makers, managers and practitioners (Petticrew and Roberts 2006: 7). The sheer volume of evidence increases daily as more research reports are published adding to the volumes of journals and books, and these are joined by different types of unpublished and non-research material. The expansion of information technologies, notably the World Wide Web, has exacerbated this information overload rather than controlling it. Web-based sources of information about research allow access to more evidence, but with the additional problem of uncertainty as to the relevance, reliability and quality of material from such sources.

Dealing with the information mountain – reviewing the evidence

First generation ‘traditional’ literature reviews

Literature reviews are a well-established method in the social sciences for trying to bring together evidence from research. These reviews have been referred to as narrative reviews (Dixon-Woods et al. 2004; Pawson and Bellaby 2006) although this label can be confusing as some narrative approaches to synthesis are far more methodologically rigorous than traditional literature reviews and can form part of systematic review processes. In the past, literature reviews were typically conducted by acknowledged ‘experts’ in the field, who would collect together individual studies they were familiar with, and attempt to make sense of the cumulative evidence by summarising and interpreting that literature. A literature review remains an essential precursor to research, a way of locating where previous research has reached, identifying the gaps and where to look next. This approach to literature reviewing still forms the basis of most doctoral theses and is often the format for the opening chapter of research reports. Many reviews written in this way make excellent reading and may well be regarded as providing sufficient of an overview to inform future research and decision-making.

However, what might be termed ‘first generation’ or traditional literature reviews paid little, if any, attention to assessing the methodological quality of
the studies included or to searching systematically for all potentially relevant evidence. Rather they tended to focus on selectively gathering and reviewing evidence that provided both context and substance to the authors’ overall arguments. Authors could be highly subjective, biased or plain ignorant in their choice of literature and interpretation. Few were critical of the research they included, and poorly designed and executed studies might be reviewed alongside high quality ones with little consideration given to how study quality might affect the results and the weight to be given to different sources of evidence. First generation reviews have been rightly criticised for a lack of transparency in the methods adopted for gathering and reviewing literature and for a lack of attention to the important issue of bias. However, as the review of research on anti-social behaviour by Rutter and colleagues (1998) illustrates (Box 1.1) it would be a mistake to assume that all such literature reviews necessarily lack sophistication in the review process or that they cannot contribute to the production of new knowledge or theory.

<table>
<thead>
<tr>
<th>Box 1.1</th>
<th>A critical appraisal of a first generation literature review: research on anti-social behaviour (Rutter et al. 1998)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Telling a convincing story</strong></td>
<td>The authorial voice is central to the power of this literature review helping to draw the reader through a complex array of material, painting a clear and coherent picture of the literature being reviewed and establishing a tone of authority and trustworthiness. The authors have essentially selected studies to provide both context for and substance to the story they are developing.</td>
</tr>
<tr>
<td><strong>2. Developing a theoretical framework</strong></td>
<td>The authors provide detailed articulation of conceptual frameworks and hypotheses developed iteratively as the review proceeds and in turn shaping the review and analysis. The review provides ‘full-colour’ illustrations of what the expert-authors judge to be key studies, key findings, key contradictions, and key gaps.</td>
</tr>
<tr>
<td><strong>3. Transparency, quality appraisal and bias</strong></td>
<td>This review includes little if any discussion about how the studies were identified and/or selected, but the approach would appear to have involved sampling for relevance rather than a systematic or exhaustive search. Similarly, there is no explicit discussion of study quality although an implicit hierarchy of evidence informs the review with longitudinal survey research presented as the ‘gold standard’.</td>
</tr>
</tbody>
</table>
Second generation reviews

In recent years a ‘second generation’ of literature reviews has developed. These adopt to varying degrees the tenets of the systematic review (described below), notably following formal and transparent review processes; using explicit approaches to the identification and selection of evidence; and attending to the methodological quality of the studies included. Whilst they do not usually transform the evidence included into a common rubric, they explicitly seek to move beyond a thin description of the evidence to produce higher order syntheses resulting in the production of new knowledge and/or theory. An example of this type of literature review is the work by Greenhalgh et al. (2005) on organisational innovation, illustrated in Box 1.2. This review provides an account of how, why and in what sequence a field of research has unfolded, enabling the reader to see how explanations (theories) and empirical findings have intertwined and changed one another through time.
Exploratory searching and mapping of potentially relevant literature in a range of different areas by members of the research team from different disciplines (e.g. ‘getting research into practice’, ‘diffusion of innovations in organisations’, etc.) because of the size of the potentially relevant literature.

Feedback to other team members of examples of ‘landmark’ primary research papers in each of the different areas identified, which demonstrated the very wide range of different designs and theoretical orientations of studies undertaken over a forty-year period.

Revision of the original review question to: ‘What research has been done that would help us understand the spread of innovations in organisations; what questions did the various research teams ask; and what did they find?’

Exploratory mapping showed that research on innovations tended to wax and wane within disciplines over time following ‘break-through’ pieces of research. The team decided to organise the review around ‘research traditions’ which used different conceptual models of the diffusion of innovations and, within each, to identify how earlier work had led to later work using citation and reference tracking.

No single set of quality criteria, rather each study was judged according to the quality criteria relevant to the research tradition to which it belonged.

Initial presentation of findings on the spread of innovations in terms of 13 different ‘research traditions’ (e.g. those of rural sociology, evidence-based medicine and guideline implementation, communication studies and knowledge utilisation) each with its own landmark studies, core scientific paradigm (conceptual and methodological) and style of presentation which evolved over time (in contrast to an attempt at a comprehensive review).

Synthesis phase in which the findings from each research tradition were related to seven key dimensions of the spread and sustainability of innovations which crossed research traditions (e.g. innovations, adopters and adoption, communication and influence, organisational context, etc.).

Review able to identify many evidence-based options for spreading good ideas which could be tried in health services.
The strengths and limitations of literature reviews

Dixon-Woods and colleagues (2004) have argued that literature, or ‘narrative’ reviews as they called them, are flexible, allowing for the inclusion of different types of evidence – qualitative and quantitative, research and non-research. This flexibility and ease of handling of a wide range of evidence means that such approaches are likely to remain an important tool for policy- and management-relevant reviews. Although the majority of these types of review still do not follow standardised procedures, as the example in Figure 2.2 illustrates, an increasing number are tending towards greater formality and explicitness in the drive for increased transparency and rigour in evidence review. At the very least, this entails paying attention to the methodological quality of the studies reviewed, and to wider issues of validity, such as the adequacy of findings (see Chapter 2 for a more detailed consideration of quality appraisal).

Both the limitations and the conceptual and analytical sophistication that can be achieved by first generation literature reviews were illustrated by the review of anti-social behaviour research undertaken by Rutter and colleagues (1998; see Box 1.1). However, as Pawson and Bellaby (2006) have argued it would be wrong to assume that the lack of standardised methods inevitably means that there is no logic to the methods used in traditional literature reviews. They suggest that literature reviews, for example those focusing on whether particular programmes or interventions work, are based on a relatively complex ‘configurational approach to causality’.

According to this perspective narrative reviews of evidence on the effectiveness of programmes or interventions are built on the assumption that positive outcomes will result from the combination of a series of program/intervention attributes . . . Interventions work, it is considered, because of the compatibility of target group, setting,
program stratagem, program content, implementation details, stake-holder alliances and so on. When using this framework it is the entire ‘recipe’ that makes the difference. All of these ingredients, along with information on outcomes and on the methodology employed in the original evaluations constitute the ontology (or the how), of narrative review.

(Pawson and Bellaby 2006: 85)

From this perspective, the task of the reviewer is to identify studies that provide the richest description of the significant properties of a particular programme or intervention. Literature reviews can be used to identify examples of ‘good practice’ or ‘best buys’ based on a judgment of the ‘fit’ between an intervention or programme and the critical success factors the review has identified.

There are practical challenges associated with traditional literature review. Their very flexibility means that the number of studies and other sources that can potentially be included could become unmanageable as could the amount of information to be extracted from studies. Additionally, diversity in the type of research makes the appraisal of study quality difficult, presents particular problems for the extraction of data in a common format and makes it hard to weight different types of evidence.

Perhaps the most important criticism of literature reviews is the potential for bias and hence for unreliable conclusions to be drawn. Shadish and colleagues (1991) suggest that in order to claim generalisability literature reviews have to demonstrate ‘proximal similarity’. This process involves selecting a feasible number of studies to review rather than attempting to be comprehensive, and choosing a manageable number of programme characteristics to explore in detail from what would certainly be a much larger number. Pawson and Bellaby question the logic of ‘proximal similarity’, suggesting that in this process some studies and factors will be privileged over others and that this introduces a whole range of biases – from those associated with publications to those arising from the personal orientation and interests of the reviewer. However, the extent to which this problem is unique to literature reviews should not be exaggerated. Recent debates in quantitative meta-analysis of trials of effectiveness have often turned on which studies to include and how much weight to give to each (Mayor 2001; Olsen 2001).

**Systematic reviews of effectiveness**

Systematic reviews of health care effectiveness developed out of a need to review health research evidence and the desire for an explicit, transparent, reproducible method for this. The systematic review is a method used to summarise, appraise and synthesize high quality evidence on effectiveness. A review is generally described as systematic if it has these features:
In addition, it is increasingly expected that systematic reviews will be regularly updated. As has already been noted, some of these features designed to ensure that the review is systematic and avoids bias are also found in ‘second generation’ literature reviews. However, an important characteristic of systematic reviews (see Cochrane reviews, below) is that the process followed is largely linear and preordained, for example, the review question or the data included cannot be altered on the basis of emerging findings or analysis of the evidence. Some approaches to synthesis, and to reviewing more generally, have a less linear, more iterative process and this can be the source of some tension between the different approaches to reviewing.

In the world of health research, systematic reviews are closely allied to the evidence-based medicine movement which has, more recently, extended to encompass policy and professional practice more broadly. In 1972, a lecture by the epidemiologist Archie Cochrane (Cochrane 1972) was published which berated medical practice as ineffective and at worst harmful to patients. He advocated the use of randomised controlled trial (RCT) methods to test the effects of interventions and argued that the findings from such trials should form the basis of decisions about practice, and perhaps more importantly decisions about health care priorities and funding: only those treatments shown to be effective should be funded. The RCT has become widely accepted as the best method for evaluating effectiveness within Cochrane systematic reviews. Cochrane also attacked the medical profession for failing adequately to summarise or collect the findings of existing RCTs. This spawned various efforts by individuals and organisations to collect and review RCTs of health care interventions. One of the earliest attempts to address the lack of knowledge about effectiveness was led by Chalmers and colleagues (1989) who spent a decade or more first producing a database of effective interventions in pregnancy and childbirth, and then testing methods for synthesizing the results from multiple studies. Efforts to develop systematic reviews have led to the creation of worldwide organisations such as the Cochrane and Campbell Collaborations which collect and review research evidence on the effectiveness of health and other social policy interventions (see Box 1.3) and to the development of resources such as the Cochrane Library of Systematic Reviews.

Interest in reviewing evidence grew during the late 1970s particularly in the USA in the field of education and by the 1980s the technique of
meta-analysis was established for aggregating statistical findings (Glass et al. 1981; Hunter and Schmidt 2004). Meta-analysis is a statistical technique used to summarise effect size from similar quantitative studies by pooling their results. Pooling the results of quantitative studies increases the statistical power of the analysis (e.g. of the relative effectiveness of two different treatments for the same condition) to detect even small effects: the more studies

**Box 1.3 The Cochrane and Campbell collaborations, and the UK NHS Centre for Reviews and Dissemination**

The UK Cochrane Centre was established at the end of 1992, by the National Health Service Research and Development Programme. ‘to facilitate and co-ordinate the preparation and maintenance of systematic reviews of randomized controlled trials of health care’. The UK Cochrane Centre is now one of twelve Cochrane Centres around the world which provide the infrastructure for co-ordinating The Cochrane Collaboration. The Cochrane Collaboration was founded in 1993 and named after the British epidemiologist, Archie Cochrane. The Cochrane Collaboration is an international non-profit and independent organisation, dedicated to making up-to-date, accurate information about the effects of health care readily available worldwide. It produces and disseminates systematic reviews of health care interventions and promotes the search for evidence in the form of clinical trials and other studies of interventions.

http://www.cochrane.co.uk/en/collaboration.htm accessed 27/10/06

The more recent international Campbell Collaboration (established in 1999) prepares, maintains and disseminates systematic reviews of studies of interventions in areas of social policy other than health care, such as education, welfare and policing and prisons. It acquires and promotes access to information about trials of interventions and builds summaries and electronic brochures of reviews and reports of trials for policy makers, practitioners, researchers and the public.

http://www.campbellcollaboration.org/ accessed 27/10/06

The UK Centre for Reviews and Dissemination (CRD) was established in January 1994, with funding from the English Department of Health and aims to provide research-based information about the effects of interventions used in health and social care. It helps to promote the use of research-based knowledge, by undertaking systematic reviews on selected topics and scoping reviews of the research literature; maintaining databases of reviews; and disseminating guidance about best practice methods for systemic reviews including, of late, reviewing qualitative research.

http://www.york.ac.uk/inst/crd/aboutcrd.htm accessed 27/10/06
that can be pooled, the greater the precision of the estimate of effect. Meta-analysis also enabled analysts to investigate the reasons for statistical variations between studies to see to what extent these were due to chance. Powered by such methods, systematic reviewing proliferated in the 1980s and 1990s and was increasingly seen as a resource not just for evidence-based medicine, but for the wider realms of public policy- and decision-making (Davies et al. 2000).

Systematic reviews of effectiveness are based on the assumption that there is a hierarchy of evidence (Box 1.4) – that some forms of evidence are ‘better’, or more valuable than others. From this viewpoint, systematic reviews of effectiveness are seen as carrying greater weight than a single study. This has been the subject of a great deal of debate and controversy with some commentators arguing that it is inappropriate to extend the notion of evidence hierarchies beyond questions of effectiveness.

**Moving beyond effectiveness reviews**

It is increasingly recognised that the hierarchy of evidence applies principally to effectiveness reviews and is less appropriate for wider reviewing activities. Whilst the more complex questions asked by policy- and decision-makers often include a concern with effectiveness (where the hierarchy may be relevant) they move well beyond this focus. Moreover there is a growing consensus (see for example Popay and Williams 1998; Petticrew and Roberts 2005) that methods for evidence review and synthesis have to be tailored to review questions. For the more complex questions facing policy- and decision-makers (see Box 1.5) a myriad of other forms of evidence in the widest sense – including qualitative research, non-trial based quantitative research, views of stakeholders and expert panels – will potentially be relevant. So, for example, research in Canada shows that health authority decision-makers take a broad and pragmatic view of what constitutes relevant evidence for

<table>
<thead>
<tr>
<th>Box 1.4</th>
<th>The evidence hierarchy for effectiveness reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Systematic reviews and meta-analysis</td>
<td></td>
</tr>
<tr>
<td>• RCTs</td>
<td></td>
</tr>
<tr>
<td>• Quasi experimental designs – cohort studies, case-control studies</td>
<td></td>
</tr>
<tr>
<td>• Surveys</td>
<td></td>
</tr>
<tr>
<td>• Case reports</td>
<td></td>
</tr>
<tr>
<td>• Qualitative methods</td>
<td></td>
</tr>
<tr>
<td>• Anecdote/expert or user opinion</td>
<td></td>
</tr>
</tbody>
</table>
priority-setting, including non-randomised quantitative studies, qualitative research, expert opinion and other more subjective sources such as anecdotal reports (Mitton et al. 2004).

Many systematic reviews, including those following the Cochrane model attempt to answer the wider questions that policy-makers typically ask. As a result, there is greater recognition of the contribution of diverse kinds of evidence, notably qualitative research evidence. This is partly linked to the growing acceptability and recognised utility of these methods in primary health-related research where qualitative methods are increasingly used alongside quantitative methods on the grounds that the two sets of data can be complementary or interactive, with the findings of one prompting questions and lines of analysis for the other (O’Cathain and Thomas 2006). The same logic potentially applies to reviews but is less apparent perhaps because methods for incorporating qualitative research and other evidence in systematic reviews are relatively underdeveloped and present a major methodological and practical developmental challenge (Dixon-Woods et al. 2001; Harden et al. 2004).

**Different objectives of reviews – ‘knowledge support’ versus ‘decision support’**

A lot of the debate about the appropriate scope of evidence that is legitimate to include in systematic reviews arises because of confusion about the purpose of different reviews. As well as recognising that the evidence and methods used
for a review should relate to the questions to be answered, it is worth understanding the different objectives of reviews. It is particularly helpful to distinguish between reviews that aim to provide ‘knowledge support’ and those that attempt to provide ‘decision support’ (Dowie 2001; Mays et al. 2005). It is important to establish which of these objectives a review seeks to meet at the outset. A review for knowledge support will be confined to summarising and synthesizing research evidence whereas the decision support function is served by a review that also includes some or all of the remaining analytical tasks required to reach a decision in a particular context.

A knowledge support review could be undertaken to answer the question ‘What does the qualitative and quantitative research evidence tell us about adherence to medication by patients with asthma?’ This might entail the following supplementary questions (Greenhalgh et al. 2004a):

- What do we know about the rates of adherence to different medications for this disease?
- Why do patients take or not take their medicines?
- Is there anything ‘special’ or unusual about asthma that might make adherence different in some way for these patients?
- Are there any population characteristics to consider (e.g. age, socioeconomic class or racial differences) that could affect adherence?
- What techniques have been tried to increase adherence to asthma medicines?
- Is adherence an acceptable or shared goal for patients and their health care providers?

These types of question can be quite broad. By contrast the kinds of question suited to decision support type reviews are likely to be much more focused and specific. A decision support review question might be ‘Should the UK NHS invest more in cognitive behavioural therapy (CBT) for people with bipolar disorders?’ Such a review would need to go beyond effectiveness questions to ask, for example:

- Does CBT work?
- Do other treatments also work?
- Which treatment works best?
- Is CBT cost-effective compared with other treatments?
- What level of improvement is shown in patients undergoing CBT and is it sufficient to alter people’s lives (e.g. so that they can return to work or normal activities)?
- How much is currently spent on CBT and how is it spent in the UK?
- How much of the existing spending could be better spent?
- What are the resource implications of investing in CBT for all the
patients who currently do not receive it but who could benefit cost-effectively?
- How acceptable would CBT be to the people at whom it would be targeted?
- What are the competing priorities in mental health care and more widely?
- What are the supply issues related to such investment (i.e. staff and facilities)?
- How feasible is it to increase CBT capacity?

The distinction between knowledge and decision support has major implications for the choice of approach to reviewing and within this to the method for synthesizing evidence. Systematic reviews designed to support policy decision-making may use more than one synthesis method within the same review. For example, a systematic review may include quantitative evidence of the effectiveness of different interventions using statistical meta-analysis, alongside a qualitative (and non-statistical) synthesis of the qualitative research evidence of their acceptability. The function of the review will influence the research questions and is likely to lead to a differing emphasis on qualitative and quantitative evidence; the more a review aims to contribute directly to a specific decision, the more it is likely to have to include non-research evidence as well as methods of modelling and simulation (see below).

**Synthesis**

Evidence synthesis embodies the idea, contained in the dictionary definition (Box 1.6) of making a new whole out of the parts: individual studies or pieces of evidence are somehow combined to produce a coherent whole, in the form of an argument, theory or conclusions. Synthesis is thus a distinct element in the review process. When the process of undertaking a systematic review is represented in linear form, synthesis is viewed as an activity at or close to the end of the review process. It occurs once the evidence has been accumulated and the data of interest extracted (see Box 1.7).

<table>
<thead>
<tr>
<th>Box 1.6</th>
<th>A definition of synthesis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Synthesis</strong> <em>noun.</em> Building up; putting together; making a whole out of parts; the combination of separate elements of thought into a whole; reasoning from principles to a conclusion (opp. to analysis). Chambers Dictionary, 1992.</td>
<td></td>
</tr>
</tbody>
</table>
Hammersley distinguishes three meanings for the word synthesis in the context of systematic reviews. First it can mean ‘aggregative’ – that is focused on the cumulation and generalisation of evidence (as in the Cochrane and Campbell collaborations); second it can mean comparative or ‘replicative’ – that is focused on the extent to which different sources of evidence reinforce one another by comparison between sources); and third it can mean that the review focused on developing ‘theory’ or explanations (Hammersley 2006: 240–1). Systematic reviews are often seen as aggregative by researchers and policy-makers/managers, but the developmental role of reviews may be equally important, for example, in generating early policy ideas or theories that can inform the development of new interventions with a plausible change mechanism underpinning them. The developmental role of reviews is particularly associated with approaches to synthesis such as realist reviews (see Chapter 5).
and meta-ethnography (see Chapter 4) but all forms of synthesis could make this type of contribution if reviewers designed these aims into the review.

There are a number of different approaches to synthesis that have potential to be applied to varied research evidence and may also have potential to be used for the review of non-research sources of evidence. However, practical experience and good examples of specific methods are limited, in some cases to a single exemplar. Most of the methods for synthesis discussed in this book were developed for primary research data analysis rather than for synthesizing research evidence. Of the methods available, most were developed either for synthesizing qualitative or quantitative research evidence rather than for combining the two in a single synthesis. However it is hoped that presenting and discussing the potential application of these methods in this book will encourage further attempts to synthesize qualitative and quantitative evidence, and persuade reviewers to explore ways of combining and integrating different kinds of evidence to inform policy- and decision-making and to contribute to the further development of methods in this field.

Noblit and Hare (1988) make a useful distinction between ‘interpretive’ and ‘integrative’ forms of synthesis which relates to the nature of the synthesis rather than the use to which it will be put (i.e. knowledge or decision support). Interpretive synthesis combines evidence by seeking to develop new concepts and theories (interpretations). Integrative synthesis can be seen as focusing on collecting and summarising data, (i.e. ‘knowledge accumulation’) although these approaches may also produce new knowledge through a ‘meta-analysis’ for example that provides a new estimate of the direction and size of effect. Borrowing from Hammersley’s typology of synthesis Dixon-Woods et al. (2006a: 36) have proposed substituting the term aggregative synthesis for integrative synthesis. ‘Aggregative synthesis’ is a term that could be applied to many literature reviews and systematic reviews, whereas the emphasis on interpretation in interpretive synthesis has meant that these approaches have largely been associated with the synthesis of qualitative research findings (see Chapter 4).

Is it feasible to synthesize disparate evidence?

There is considerable controversy about the legitimacy and/or feasibility of combining the findings of different types of evidence, notably evidence from different kinds of research. Within the qualitative research community, in particular, it is suggested that attempts at aggregative synthesis destroy the integrity of individual studies thereby producing meaningless findings (Sandelowski et al. 1997). This critique is informed by a relatively extreme ‘relativist’ position which argues that qualitative research offers multiple ‘truths’ or realities such that each study represents a unique, personalised view...
that cannot be replicated, added to another or transferred. This argument is further complicated when a synthesis attempts to combine findings across different methods or studies informed by different theories of knowledge: relativists, for example, would suggest that differences in theory and method are fundamental and militate against any integration of research. However, a ‘subtle realist’ view (Hammersley 1992) would suggest that while there may well be multiple descriptions or explanations of phenomena and ways of knowing, these ultimately relate to some underlying reality or truth. From this perspective, synthesis is accepted as potentially promoting a greater understanding than any single study could achieve.

There are those who contend that the divide between the qualitative and quantitative research paradigms is far greater than the differences between different types of qualitative research, and that this is an irremovable barrier to the synthesis of qualitative and quantitative evidence. However, to some extent this divide between qualitative and quantitative research is being bridged in primary studies as many researchers now engage productively in multi-method research, utilising both quantitative and qualitative methods (O’Cathain and Thomas 2006). Synthesis may be seen as a logical extension of this combined approach. Just as in mixed method primary research, qualitative data can illuminate why a particular policy or management approach has variable impacts and suggest ways of dealing with this, and quantitative studies can indicate its relative effectiveness overall: both have a contribution to make to understanding the processes shaping implementation of interventions/programmes.

The feasibility of synthesizing non-research evidence has been little explored but similar issues arise regarding the context and applicability of non-research evidence and these are overlaid with concerns about validity and rigour. Nonetheless, these forms of evidence may well be necessary components of decision-making in messy, complex policy arenas and the question here may be less ‘is it feasible to synthesize these kinds of evidence?’ and more ‘is it legitimate to ignore them?’

Evidence synthesis can help to inform decision-making by managers and policy-makers. It can answer questions which individual studies or other pieces of evidence cannot – by integrating and interpreting complex and diverse kinds of evidence in a form which is more manageable and accessible. Used for knowledge support synthesis can inform – identifying gaps where further research is needed, locating controversy and mapping the terrain. In a decision support mode, synthesis can be used to help decide between competing priorities or alternatives.

The next chapter looks in more detail at the process of reviewing evidence systematically to provide the foundation for synthesis.