Producing a Systematic Review

David Denyer and David Tranfield

INTRODUCTION

The aim of this chapter is to provide guidance to scholars, practitioners, and policy makers who are engaged in producing, commissioning, or using reviews of research evidence in the field of management and organization studies. In contrast with many other natural and social science fields, inexperienced researchers, particularly doctoral students in management and organization studies, often receive relatively little training in conducting research reviews. Compared to the wealth of texts on philosophical approaches to social science research and methods for empirical investigation, there are few instructional texts on literature reviewing. This is surprising given the critical role that literature reviews play in doctoral theses and journal publications, and the potential role that they could play in creating and building bodies of knowledge and informing policy and practice (Tranfield et al., 2003).

We argue that researchers in management and organization studies have a significant opportunity to evaluate and learn from many other fields that have developed an evidence-based approach using systematic review as a key technique. However, different academic fields have idiosyncratic characteristics and as a result require the development of bespoke approaches specifically tailored to serve their particular purposes, forms, and applications.

Systematic review is a specific methodology that locates existing studies, selects and evaluates contributions, analyses and synthesizes data, and reports the evidence in such a way that allows reasonably clear conclusions to be reached about what is and is not known. A systematic review should not be regarded as a literature review in the traditional sense, but as a self-contained research project in itself that explores a clearly specified question, usually derived from a policy or practice problem, using existing studies. Additionally, systematic review also differs from other review methods because of its distinct and exacting principles. For example, in a systematic review, the researcher is required to set prespecified relevance and quality criteria for the selection/inclusion of studies and to make such criteria transparent to readers. Extensive searches are conducted to incorporate both published and unpublished studies. In terms of outcome, where studies provide consistent results, systematic reviews
might be expected to provide solid and dependable evidence that is robust and has potential for transfer across different contexts. On the other hand, if the review identifies knowledge gaps or incongruent findings, then this signifies a research need and raises questions for future research.

Major improvements in review techniques over the last fifteen years or so have helped to raise the profile of systematic review as an important research methodology. Applications of systematic review in medical science have led the way and are often considered to reflect the most advanced thinking. A 'standard' approach to systematic review was developed initially in the medical field by the Cochrane Collaboration (2008), and has been followed later by other consortia dedicated to commissioning and disseminating systematic reviews.

A Systematic review has been argued to bring replicable, scientific, and transparent approach, which seeks to minimize bias (NHS Centre for Reviews and Dissemination, 2001) and requires reviewers to summarize all existing information about a phenomenon in a thorough and unbiased manner. More widely, systematic reviews have been argued also to have value in collating and synthesizing existing evidence across a wide range of settings (including the social sciences) and empirical methods (Petticrew, 2001).

On the other hand the use of systematic review, particularly its traditional form, has been criticized when applied to the social sciences (Hammersley, 2001). Usually, these criticisms have centred on uncritical mimetic application encouraging the simple transfer of the medical model of systematic review, exemplified by the Cochrane Collaboration (2008), into other fields. Critics have placed four sets of arguments at the centre of this debate concerning:

- The extent to which systematic review methods can or should be proceduralized and the extent to which bias can be reduced or mitigated.
- The validity and reliability of synthesized evidence produced by systematic reviews, and how such evidence is used subsequently to inform policy, practice, and future research.

Other critics have suggested that evidence-based management is overtly managerialist (Learmonth and Harding, 2006) and fails to take into account situated judgement and ethics (Morrell, 2008). They also express a concern that evidence-based management and systematic review privileges certain types of research and other forms, such as critical theory, are precluded.

We appreciate many of these concerns and challenges and take the view that it is unwise to adopt, in an uncritical fashion, that which has proved useful in other fields. Consequently, our purpose in this Chapter is to explain how and in what ways the notion of evidence-based management might benefit the field through improved quality and rigour, by defining a bespoke and fit for purpose format for review. Differences in context are vital, for example, in medicine, there has been a general consensus regarding what constitutes an appropriate methodology to be used for evaluating evidence from primary studies to be included (Evans, 2003). In fields where the value of quantitative, qualitative, and theoretical contributions are equally respected (such as management and organization studies), reviewers are likely to need more inclusive and comprehensive quality criteria and novel approaches to research synthesis. Such inherent epistemological diversity necessitates an alternative to 'Cochrane-style' systematic reviews. Therefore, we argue for the development of a bespoke and fit for purpose methodology, which can cope with the variety and richness of research designs, purposes, and potential end uses of management and organization studies reviews.

Despite these differences our point of departure is that systematic review methodology has proved demonstrably worthwhile in
many natural and social science fields, and that would be wise to consider the lessons researchers in management and organization studies already learned. Consequently, this chapter first outlines the principles of systematic review developed elsewhere and evaluates their relevance for the field of management and organization studies. We discuss both the origins of systematic review and its roots in evidence-based policy and practice and also examine the methodological challenges to developing an evidence base in management and organization studies. We note that some of the key principles of systematic review that are manifest in fields such as medicine, often remain elusive and perhaps undesirable for the field of management and organization studies. Second, and on the basis of this considered reflection, we offer a revised and (to an extent) alternative framework and set of principles for developing transparent, inclusive, explanatory and heuristic reviews for use in our field. Finally, we explore how the output from systematic reviews might be used to inform policy, practice and further research. In doing so, we argue that the methods of review should be selected according to the purpose and the nature and characteristics of the field of study (e.g. degree of fragmentation, plurality of methods, etc) and the nature of the evidence available. We argue the case for combining the many facets of systematic review with aspects of realist synthesis (Pawson, 2006). Harnessing the strength of both of these approaches offers an important methodology for management and organization studies that has the potential to inform policy, practice, and future research.

The origins and principles of systematic review

The idea of reviewing existing research studies for the purpose of informing policy, practice, and research is not new. However, over recent years, there has been a proliferation of formal approaches and systematic methods for locating, selecting, appraising, synthesizing, and reporting evidence. This can be seen as a response to the increasing demands to organize knowledge into a format that is rigorous and reliable as well as make a difference to practice. Evidence-based practice requires decision making and action to be informed by the best and robust evidence produced by high-quality research studies. In addition to the traditional approach of conducting new empirical studies, many natural and social science fields have turned to reviews of existing primary studies reported in the literature to address important and pressing questions. In this demanding context, systematic review increasingly has become a standard method for locating, selecting, and appraising research and transferring the synthesized findings not only to researchers and academics but also to practitioners and policymakers in a digestible format to inform action.

Evidence-based policy and practice (EBPP) constitutes a movement that can be traced to the early 1970s (Cochrane, 1972). EBPP in medicine can be defined as the ‘conscientious, explicit and judicious use of current best evidence in making decisions’ about the care of individual patients (Sackett et al., 1996: 1). Since the early 1980s, researchers in health care have been engaged in a programme of systematic reviews on the effectiveness of interventions. The Cochrane Collaboration was formed in 1992 to gather, evaluate, and disseminate research evidence and developed a sophisticated and substantial set of guidelines for conducting systematic reviews. Since then, researchers have developed and refined the systematic review methodology with the aim of improving both the rigour and reliability of the review process, and organizing knowledge into a usable format. The results of systematic reviews have become well established as legitimate evidence on which to base policy and practice. Websites, such as the Cochrane database (2008b), now offer clinicians a practical solution to the problem of staying up to date with the latest research.

Given the relative success of evidence-based approaches in medicine, the core
principles, processes, and practices of systematic review have been adopted elsewhere. As Petticrew (2001: 99) argues:

...systematic review is an efficient technique for hypothesis testing, for summarizing the results of existing studies, and for assessing the consistency among previous studies; these tasks are clearly not unique to medicine.

Following this line of thinking, the Campbell Collaboration was formed in 1999 to establish the systematic review process in the broader public policy arena (Campbell Collaboration, 2008). The Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) conducts systematic reviews in many fields, including education and social care, and has developed an approach for addressing a broad range of review questions and incorporating both quantitative and qualitative evidence.

In general terms most fields using systematic review prescribe that studies are expected to conform to a set of principles as well as adopting specific methods to identify, select, and critically appraise relevant primary research, before it is appropriate to extract and analyse data from the studies that are included in the review (NHS Centre for Reviews and Dissemination, 2001).

In the following section, we outline the four core principles traditionally applied to a standard systematic review template. Reviews are expected to be: replicable, exclusive, aggregative, and algorithmic. We then contrast these with a suggested set of principles for systematic reviews in management and organization studies.

**Replicable**

A first key principle suggests that the general methodological characteristics of systematic review reflect an epistemology stipulating that knowledge should be acquired through a ‘scientific’, objective using an unbiased process. Systematic reviews, particularly those associated with the Cochrane Collaboration, attempt to avoid or to mitigate bias. As such, ‘systematic reviews have been developed to synthesize research according to an explicit and reproducible methodology’ (Greenhalgh, 1997). As in any scientific endeavour, the methods to be used are required to be established beforehand (Light and Pillemer, 1984). Consequently, prior to a systematic review being undertaken, a protocol is produced to detail precisely how the review will be conducted. The protocol is analogous to and as crucial as a research design for an empirical study (Antman et al., 1992; Cook et al., 1997). In medicine and other fields, review protocols must be developed and approved before the systematic review can commence. Reviewers are encouraged to publish their draft protocols (sometimes on these websites) allowing interested parties to provide feedback. While the intention should be that a review will adhere to the published protocol, it is acknowledged that the review protocol may need to be changed during the course of the review. If modifications are made to the protocol, they must be documented explicitly and explained (Higgins and Green, 2006).

Systematic reviews have a clearly defined methods section with each step of the systematic review rigorously reported. Justifications are given for all decisions taken by the reviewer. The aim is to mitigate the preformed opinions that can bias his/her assessments of both the relevance and validity of articles in a review (Cooper and Hedges, 1994). Only studies that meet the inclusion and exclusion criteria are included in the review. The reviewer also needs to discuss whether or not all relevant studies were identified, whether or not all relevant data could be obtained, and whether or not the methods used (for example, searching, study selection, data extraction, and analysis and reporting) could have introduced bias. The aim is for systematic reviews to:

...bring the same level of rigour to reviewing research evidence as should be used in producing that research evidence in the first place (Davies and Crombie, 2001: 1).

The use of explicit, systematic methods in reviews has been argued both to help produce
more reliable results upon which to draw conclusions and make decisions (Antman et al., 1992). Cooper and Hedges (1994: ix) also argue that the ‘intended result is a research synthesis that can be replicated by others’.

However, within a review of social science literature, it is generally recognized that professional judgment and interpretation play an important role and cannot be eliminated or replaced by proceduralization. Hammersley (2001: 6) argues that:

...insisting that people follow a set of procedures rigidly can be a hindrance. Literally, it may rule out any reflection on issues that actually need attention, encouraging a rather pragmatic orientation. And this is especially likely where the time available for producing a review is very limited. As Polanyi points out, science, like many other activities, relies on skilled judgement, the capacity for which is built up through experience; and skilled judgement means taking account of particulars, and will often involve apparently breaking the rules.

**Exclusive**

A second key principle requires that if systematic reviews are to inform policy, practice, and future research, they should synthesize the best evidence available (Slavin, 1986). In systematic reviews, the validity of a study is governed by the extent to which its design and conduct are likely to prevent systematic errors, or bias (Moher et al., 1995). A review:

...should always focus on the best available evidence: that is, studies to be included in a review should use methods that provide the least biased answer to the question asked (Counsell, 1997: 380).

The Cochrane Reviewers’ Handbook (Higgins and Green, 2006) and the NHS Centre for Reviews and Dissemination guidelines (2001) both suggest that quality relates to the extent to which the study minimizes bias and maximizes internal and external validity. Relevant studies are critically evaluated to assess their quality using checklists that are included in the protocol along with details of how the checklist was used in the study.

**Table 39.1 The hierarchy of evidence in medical science (Davies, Nutley and Smith, 1999: 11)**

<table>
<thead>
<tr>
<th>I</th>
<th>Systematic review and meta-analysis of two or more double blind randomized controlled trials.</th>
</tr>
</thead>
<tbody>
<tr>
<td>I-1</td>
<td>One or more large double-blind randomized controlled trials.</td>
</tr>
<tr>
<td>I-2</td>
<td>One or more well-conducted cohort studies.</td>
</tr>
<tr>
<td>II-1</td>
<td>One or more well-conducted case-control studies.</td>
</tr>
<tr>
<td>II-2</td>
<td>A dramatic uncontrolled experiment.</td>
</tr>
<tr>
<td>III</td>
<td>Expert committee sitting in review; peer leader opinion.</td>
</tr>
<tr>
<td>IV</td>
<td>Personal experience.</td>
</tr>
</tbody>
</table>

Typically, systematic reviews in medicine adopt a hierarchy of evidence (Table 39.1) that ranks research designs according to their internal and external validity. Studies are selected on the basis of relevance for inclusion (ideally assessed by two reviewers). In general, randomized controlled trials are regarded as the ‘gold standard’ for judging ‘what works’ (Evans, 2003). For example, Oakley (2000: 318, quoting Macdonald, 1996: 21) claims that anything except randomized controlled trials ‘is a “disservice to vulnerable people”’.

Whilst most evidence in medicine is made available in quantitative form, a number of systematic reviews include evidence from qualitative studies. Consequently, a set of quality standards for qualitative research have been developed (see, for example, Harden et al., 1999; Oliver et al., 2005; Popay et al., 1998) although the matter of what constitutes appropriate standards to apply to qualitative research still remains contested. More usually, evidence from qualitative studies is rarely included in the actual synthesis process, and is more usually descriptively summarized in the systematic review report. Whilst efforts have been made to incorporate qualitative evidence in medical science, there remains a privileging of quantitative over qualitative research (Dixon-Woods et al., 2005).

In the field of management and organization studies, there is a plurality of accepted methods and approaches. Whilst this variety is often regarded as a weakness, it also provides
diversity and richness which are a major strengths considering that:

Organizations are fuzzy, ambiguous, complex socially constructed systems that cannot be well understood from a single perspective (Denyer et al., 2008).

However, there is no doubt that there is little agreement over what represents high quality evidence across this field (Tranfield and Starkey, 1998) and certainly there is robust opposition to prioritizing one research method over another.

**Aggregative**

The third key principle suggests that alongside systematic review, statistical meta-analysis has become the preferred approach to research synthesis in many domains. A meta-analysis extracts and aggregates data from separate studies to increase the effective sample size, in order to calculate an overall effect size for an intervention. In order to perform a meta-analysis, it is important to identify whether or not results from studies are consistent one with another (i.e. homogeneous) or inconsistent (e.g. heterogeneous). Data from individual studies are extracted and tabulated in systematic review reports to display the impact of potential sources of heterogeneity, e.g., study type, study quality, and sample size.

If the data extracted from these studies meet certain requirements (the most important being a high level of homogeneity of effect measures across studies) then they can be aggregated. Pooling data provides a larger sample, increasing the possibility of detecting real effects that individual smaller studies are unable to ascertain. However, aggregating studies can also cause problems if the researcher fails to account for subtle differences in study design and study context, since it is possible to reveal minor biases as well as true effects.

Therefore, a key problem with aggregative synthesis is that primary studies are rarely entirely homogeneous. However, there are strong arguments that studies can be pooled despite subtle differences in effects or lack of correspondence in the populations or study contexts such as variation in form of implementation, measured outcomes, or contextual factors that influence observed effects (Rousseau et al., 2008). Yet, the danger remains that unless primary studies carefully report implementation processes and context, aggregative syntheses can mask the mechanisms underlying effects (Pawson, 2002). Advocates of meta-analysis (Rauch and Frese, 2006; Shadish and Sweeney, 1991) argue that heterogeneity can be dealt with by including mediators and moderator variables into statistical techniques. A moderator is an intermediate variable that links an intervention with an outcome. A moderator accounts for contextual differences, either in terms of the population or the circumstances. A host of demographic and contextual factors can be specified and accounted for in the analysis. The problem with this technique is that each of the likely mediators and moderators must be identified at the outset of the review, and it is likely that in a managerial and organizational context the list of variables could be extensive. Further, each of the studies included in the review must have comparable mediators and moderators, thus severely limiting the number of studies on any specific topic that are amenable to aggregative synthesis.

Additionally, despite advances in meta analytical techniques, many questions are not amenable to aggregative approaches. In complex systems comprised of a large variety of dynamically interacting variables such as can be found in management and organizational studies, the value of statistical synthesis adopted as a single point solution is often questionable. Many systematic reviews, even in medicine, do not contain a meta-analysis. Dixon-Woods et al. (2006) assert that alternative approaches to synthesis, including qualitative techniques, can be incorporated into systematic reviews.

**Algorithmic**

In health care, systematic reviews tend to focus on questions relating to the effectiveness
of interventions. A large number of reviews provide a comparison of two or more alternative treatments to identify which is the most effective. Systematic reviews in medical science tend to address specific questions structured according to three features; the population, the intervention, and the outcomes. A review may focus on very specific population groups by addressing factors such as the age and sex of the patients or the severity of their illness. Reviews may also specify a particular setting of interest such as whether people are living in the community or are hospitalized (NHS Centre for Reviews and Dissemination, 2001). The interventions are often the treatments given to the populations of interest, which are usually a particular therapy, prevention, diagnosis, or screening (Higgins and Green, 2006). The resulting outcomes are specific, measurable clinical and economic factors that are used for comparison. In some instances, the study designs to be included in the review are also specified. An example of a well-formatted review question is provided in the Cochrane Handbook (Higgins and Green, 2006: 62):

...whether a particular antiplatelet agent, such as aspirin, [intervention] is effective in decreasing the risks of a particular thrombotic event, stroke, [outcome] in elderly persons with a previous history of stroke.

Such algorithmic guidelines provide a procedure (a finite set of well-defined instructions) for overcoming a condition and accomplishing a defined outcome. Algorithmic rules require limited redesign for use in practice (van Aken, 2005). The problem with this approach is that it seeks to identify what works without sufficient knowledge of why an intervention works or the particular conditions of success (Pawson, 2006). To achieve richer understanding often requires practitioners to integrate evidence with their knowledge of local circumstances and the preferences of the end user. Therefore, it is usually thought wise for mature users of systematic reviews to decide how appropriate the evidence is to their particular circumstances in order to obtain optimum benefit.

Methodological challenges

A review of the literature in the field of management and organization studies is typically the first step in the process of doing scientific research, and is usually presented as either a brief introduction to an empirical study or as an extensive detailed account of a body of literature (see Academy of Management Review or the International Journal of Management Reviews for examples). Reviewing in management and organization studies is particularly challenging due to the fragmented nature of the field (Whitley, 1984) and also its transdisciplinarity (Tranfield et al., 2003). It is a field in which many separate sub fields can generate idiosyncratic questions, hypotheses, methodologies, and conclusions (Baligh et al., 1996). Subfield academic communities easily detach from identification with the whole and can often engage in their research unaware of work in associated areas:

The complicated state of MOS (Management and Organization Studies) research makes it tough to know what we know, especially as specialization spawns research communities that often don't and sometimes can't talk with each other. Organizational researchers commonly adopt positions regarding management practices and organizational decisions for which no accumulation of evidence exists, or where at least some of the evidence is at odds (Rousseau et al., 2008: 477).

Tranfield et al. (2003) argue that management research is a nascent field still developing in terms of agenda and focus. In contrast with more mature fields, such as medicine or engineering, management research does not enjoy consensus on research methods and convergence on research questions. Accordingly, researchers tend to ask and to address a steady flow of questions rather than integrate and build coherent knowledge stocks or seek further understanding of particular phenomena. Vastly increasing volumes of management research, coupled
with the potential of information technologies, has meant recently that synthesizing diverse literatures and accumulating a knowledge base serving both the research and practitioner communities, has become an increasingly topical challenge.

In addition to informing the research community, a review of existing research evidence has the potential to guide practitioners by providing ideas, information, examples, and recommendations for practice. So why have practicing managers so far rarely used systematic review to improve their performance? First, practicing managers rarely have access to academic journals or possess the skills required to interpret and appraise research evidence. Second, many reviews have been framed to answer theoretical questions for academics and do not address the problems that managers face in their work roles. The aim of academic reviews is to position an empirical study within an established body of knowledge. According to Wallace and Wray (2006) such reviews relate to one or more explicit or, more typically, implicit review questions that may be substantive (about some aspect of the social world), theoretical (about concepts, perspectives, theories, or models that relate to some aspect of the social world), or methodological (about the approach to conducting an empirical or theoretical enquiry). A traditional review seeks an understanding of the nature of the field, the major issues and debates about the topic, how the field has been structured and organized, the key theories, concepts and ideas, and the ways in which the field has been researched to date (Hart, 1998). Such reviews usually tackle knowledge problems and focus on analysis and explanation, and on problems and their causes (Denyer et al., 2008). Such research ‘criticizes everyday accounts and practices … but does not seek to transform them except in the general sense of demonstrating their incorrectness’ (Whitley, 1984: 372).

Denyer et al. (2008) argue that, if the product of a review is intended to be used by practitioners and policy makers, then it should address field problems, that is the problems faced by professionals such as clinicians, engineers, lawyers, and managers in their work roles, rather than focusing on pure knowledge problems. In so doing, they argue that systematic reviews in management and organization studies can produce knowledge of use to managers in providing insight or designing solutions to field problems.

**Revised principles for management and organization studies?**

A number of authors have argued that the field of management and organization studies might benefit from the adoption of systematic review methods (Briner, 2000; Pittaway et al., 2004; Tranfield et al., 2003). However, Tranfield et al. (2003) voice a concern that the adoption of Cochrane-style systematic review is questionable and potentially undesirable for use in the field of management and organization studies, arguing that a revised and fit for-purpose systematic review methodology might be more appropriate. However, it has remained unclear precisely which elements of the standard systematic review template should remain unaltered, which facets require amendment, what should be refuted, and what needs to be added for effective use in the field of management and organization studies. Reviews in medical science, we have noted, are based on four core principles, but all of them have limitations when applied to management and organization studies. Consequently, we now offer, four alternative principles for systematic reviews for use in management and organization studies. We suggest reviews be tested for their transparency, inclusivity, explanatory, and heuristic nature.

**Transparency**

Concerning our first key principle, we contend that the reasons for documenting the review methods is not to achieve replication or eradication of bias, as in a Cochrane-style review, but rather to aid transparency (Tranfield et al., 2003). We believe that there are three aspects to transparency in a systematic review.
First, reviewers must be open and explicit about the processes and methods employed in the review. This enables readers to determine precisely its scope and boundaries. Throughout the review, the steps undertaken need to be specified, applied, recorded, and monitored (Tranfield et al., 2003). For example, the practice of making the review methods explicit before the review commences is a constructive exercise. Producing a review protocol enables the reviewer to gain feedback on the proposed methods and highlight any obvious errors or omissions so that they can be amended or rectified. Including a methodology section in the systematic review report enables readers to determine precisely the scope and boundaries of the review. It also provides an audit trail and enables the review to be updated and appraised in the future. A systematic review protocol does not mean that the predetermined methods are set in stone. It is important that the protocol does not restrict the review and it is quite normal for reviewers to alter the protocol during the course of conducting the review. A reviewer may, for example, find a body of work that they had not foreseen or alternatively for which they may wish to alter the selection criteria as their understanding of the field develops. In such circumstances, the protocol can be used to document and justify these changes. In the final systematic review methodology section, it is customary to produce an overview of the main changes between the original and final protocols.

The second expression of transparency concerns the presentation of the findings of the review in such a way that there are clear links between the evidence found and the reviewers’ conclusions and recommendations. As noted by Slavin (1986: 17):

No review procedure can make errors impossible or eliminate any chance that reviewers’ biases will affect the conclusions drawn. However, applications of best-evidence synthesis should at least make review procedures clear to the reader and should provide the reader with enough information about the primary research on which the review is based to reach independent conclusions.

A clear audit trail is required linking the claims made by the authors of the review with the existing evidence. According to Pawson (2006: 79):

... other researchers, decision makers and other stakeholders can look behind an existing review, to assure themselves of its rigour and validity, reliability and verifiability of its findings and conclusions.

The third element to transparency necessitates the reviewer making clear the assumptions underpinning the review and engaging in a mindful questioning of a priori beliefs regarding the scope and implications of relevant research. In conducting a review, particularly problem specification, study selection and synthesis, the reviewer necessarily falls back on their values, prejudices, and beliefs. Wallace and Wray (2006: 88) argue that the ‘authors’ explicit or implicit values about some aspect of the social world, their theorizing, research methodology, and methods may affect their focus and the nature of the knowledge claims that they make’.

In particular, the reviewer’s prior knowledge of the literature may influence the review. At worst, ‘a reviewer may simply build a case in support of their personal beliefs, selectively citing studies along the way’ (Davies and Crombie, 2001: 2). As such, it is important that reviewers make explicit their value stance towards the aspect of the social world they are studying (Wallace and Wray, 2006).

**Inclusivity**

Our second key principle emphasizes that systematic reviewers in the field of organization and management are likely to encounter difficulties when appraising the quality of information sources. Authors of articles, even those published in some of the highest impact journals, sometimes fail to report on the methods of data collection and analysis in sufficient detail, rendering impossible attempts to assess study quality. Within management and organization studies, there is also little uniformity in methods of data collection
and analysis. Studies rarely address identical questions and samples vary in terms of the populations, sizes, study context, and the data reported.

Boaz and Ashby (2003: 4) suggest that rather than a hierarchy of evidence, the selection of articles can be based on the criterion ‘fit for purpose’; this, ‘helps us to get away from the technocratic preoccupation with elegant research designs and allows for notions of appropriateness to be added into the appraisal of research’. Pawson (2006) also argues that the researcher should simply ask whether or not the literature retrieved adds anything new to our understanding of the intervention. Pawson’s approach has no hierarchy of evidence. The worth of the study can only be determined through the synthesis by evaluating each study’s contribution to theory building, prioritizing the vital evidence from primary studies as the original researchers’ interpretations and explanations, not just the results. Hence, Pawson encourages the scope of the review to include a wide range of studies, research types and data forms to promote a full understanding of the phenomenon of interest.

Salipante et al. (1982) have also argued that the inclusion of a broad variety of sources in a review can compensate for researcher value judgments and uncontrolled validity threats. Heterogeneity also makes possible the investigation of contextual factors influencing the study’s design and its findings. They argue that the more heterogeneous the distribution of uncontrolled validity threats in a set of similar findings, the greater the validity of the findings from the set. Following this logic, we suggest that reviewers in management and organization studies should beware of excluding studies solely on the basis of adherence to the pursuit of absolute epistemological standards. In particular, reviewers are best advised to guard against using proxies for research quality such as the quality rating of journals as a basis for exclusion. We do, however, agree that systematic reviewers in management and organization studies devise quality checklist(s) appropriate to the subfield(s) involved, justify the reasons for inclusion/exclusion of studies, apply the criteria to all relevant studies, and communicate the warrants underpinning their claims (Wallace and Wray, 2006).

Explanatory

Our third key principle addresses the vexing question of synthesis. The evidence-based community in the UK (see for example the Evidence Network website: http://www.evidencenetwork.org/) has been working hard to develop several alternative approaches to meta-analysis as a mode of synthesis. Particularly, progress has been made in narrative reviewing (see Popay et al., 2006; Greenhalgh et al., 2005) and the use of the metaethnographic approach of Noblit and Hare (1982; see Campbell et al., 2003). Evidence syntheses using these approaches are tasked with the construction of larger narratives and more generalizable theory (Estabrooks et al., 1994; Sandelowski et al., 1997). Interpretive and explanatory synthesis extracts descriptive data and exemplars from individual studies, building them into a mosaic or map (Hammersley, 2004). The reviewer is involved in juxtaposing the evidence, which can be quantitative, qualitative, or theoretical, from one study with that of another.

Unlike aggregative synthesis which seeks to avoid or mitigate bias, interpretive and explanatory syntheses are ‘active’ and ‘creative’ methods (Pawson, 2006) that go beyond a descriptive reporting of the evidence. The process is one of conceptual innovation and reinterpretation (Campbell et al., 2003), while attempting to preserve the original study’s integrity or wholeness (Pawson, 2006). As such, the synthesis provides a feasible explanation of the study findings rather than a replicable explanation (Noblit and Hare, 1988). Therefore, synthesis involves the process of bringing the pieces from individual texts together to make a whole that should be more than the sum of the parts.

Pawson (2006) has introduced the notion of realist synthesis by using existing primary studies to build a nascent theory of how
an intervention may work. This will include ‘conjectures on the generative mechanisms that change behaviour, ideas on the contexts that influence its operation, and hunches about the different outcome patterns that ensue’ (Pawson, 2006: 73). The texts that are to be included in the synthesis are ‘regarded as case studies, whose purpose is to test, revise and refine the preliminary theory’ (Pawson, 2006: 74). The review entails the systematic organization of the data into formats that allow summary. This body of evidence is then probed, sifted, coded, and cross tabulated in numerous ways. Each relevant published article is described and discussed in terms of its contribution to the emerging theory (Pawson, 2002: 345):

…the reviewer’s basic task is to sift through the mixed fortunes of a programme attempting to discover those contexts (C+) that have produced solid and successful outcomes (O) from those contexts (C) that induce failure (O-). The review process is then repeated across a number of additional studies featuring the same underlying mechanism (M), with the aim of gathering together the various permutations of success and failure … the aim is to differentiate and accumulate evidence of positive and negative CMO configurations.

Generalization is sought not in terms of the associations among variables but in terms of the role and impact of generative mechanisms that play out in different ways over time.

**Heuristic**

The output of a systematic review in management will allude to what works, and to why or how the relation occurs and in what circumstances, but is likely to be relatively abstract and is best regarded by the practitioner as a ‘design exemplar’ (van Aken, 2005). Given the complexity of organizational settings, such outputs are likely to be heuristic. A heuristic rule may help in solving a problem, but is not guaranteed to provide a detailed solution. Outputs of systematic review in management and organization studies are likely to be rules, suggestions, guides, or prototype protocols that may be useful in making progress toward a solution of a problem rather than providing a detailed solution to a specific problem. Heuristic rules will almost always require the application of informal judgment or experience to contextualise action. Denyer et al. (2008) suggest that systematic reviews of the existing published science base in management and organization studies can develop knowledge that managers can use to design solutions to problems in their field. In management research, rather than presenting the truth, in the form of ‘valid evidence’, managers may be presented with some ‘clues/ideas’, ‘tools’, and ‘methods’ that may help to guide design for effective implementation. Deciding the degree to which the findings presented in any particular review can inform practice is therefore a matter of judgment for the practitioner.

**Applying the new principles**

**Step 1: Question formulation**

As with any research, the primary decision in preparing a review is to establish its focus (Light and Pillemar, 1984). This is done most effectively by asking clearly framed questions (Cooper and Hedges, 1994). By clearly formulating the question, criteria for primary study inclusion in the review become clear:

A good systematic review is based on a well-formulated, answerable question. The question guides the review by defining which studies will be included, what the search strategy to identify the relevant primary studies should be, and which data need to be extracted from each study. Ask a poor question and you will get a poor review’ (Counsell, 1997: 381).

In order to clarify review questions, inform the review process and improve the utilization of findings from systematic reviews, it is best to involve a broad range of stakeholders in the development of review questions and procedures (Tranfield et al., 2004). Petticrew (2001: 100) argues that,

‘potential users of systematic reviews, such as consumers and policy makers, can be involved in the process. This can help to ensure that reviews
are well focused, ask relevant questions, and are disseminated effectively to appropriate audiences'.

Advisory groups can assist with defining the broad topic area to be investigated and identifying the specific areas within that topic that would be most useful to scrutinize in depth. Involving experts is particularly important when the literature is sparse. The review panel best includes people with academic knowledge and practical expertise in the subject area such as librarians and information scientists and systematic review specialists, who may be consulted individually or as a group at key points throughout the review.

In the first instance, the review question may be free-form, written in natural language. The free-form question can then be structured into a reviewable question. In management and organization studies, it is sometimes difficult to define and measure constructs of interest. There is usually an emphasis on the requirement to understand the impact of study context on the study results. In a complex social world, it is often difficult to predict with any surety a uniform link between intervention and desirable outcome. Therefore, research in complex social systems (such as organizations) must always take into account the crucial role of context. As noted by Pawson and Tilley (1997), interventions in such systems will be affected by at least four contextual layers – the individual, interpersonal relationships, institutional setting, and wider infrastructural system. Pawson and Tilley also highlight the role of generative mechanisms, which help to explain the ephemeral relationship between the intervention and outcome. Generative mechanisms are the nucleus of explanation in realist synthesis (Pawson, 2006) and typically, for any process, there will be at least one mechanism operating (see Reed, Chapter 25, this volume). More likely, in complex social situations, there will be many mechanisms working concurrently (Denyer et al., 2008). The generative mechanism is a theory that helps to explain causal relations by ‘describing powers inherent in the system’ (Pawson, 2006: 2). Pawson (2006: 24) provides an interesting example of how organizational forms are constrained by context:

Bureaucracies organize work routines in certain ways and provide tightly specified employment functions for their workforces, but whether any of this happens depends on the availability of work, and ultimately, on the overall economic health of the nation. The efficacy of bureaucratic management is also contingent on the type of work carried out. In sectors that thrive on innovation and entrepreneurial activity, the applicability of fixed duties to fixed roles is likely to fail.

In social science fields such as organization and management studies, it is not enough to know what works but also to ascertain why or how the relation occurs and in what circumstances. Pawson’s (2006) realist approach requires the reviewer to determine (or infer) context, mechanism, and outcome configurations through comparing and contrasting interventions in different contexts. Such an approach has the advantage of including different types of information, such as case studies, so long as they provide some insight into what works, why, where, and when. Denyer et al. (2008) develop this argument using the acronym CIMO (Context, Intervention, Mechanism, and Outcome) – an acronym that can be used to specify the four critical parts of a well-built systematic review question (see Figure 39.1).

In management and organization studies, a practitioner question may be framed as follows: how can project team performance be optimized through leadership behaviours? This may be structured into a series of reviewable questions such as: under what conditions (C) does leadership style (I) influence the performance of project teams? What mechanisms operate in the influence of leadership style (I) on project team performance (O)?

In reviewing these questions, there are multiple perspectives with regard to the meaning of key constructs. Team performance, for example, may be measured in terms of project outcomes such as time, cost, and quality. However, from team members’ perspective, performance may be interpreted
<table>
<thead>
<tr>
<th>Component</th>
<th>Questions to ask</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Context (C)</strong></td>
<td>Who are the individuals of interest?</td>
</tr>
<tr>
<td></td>
<td>• Employees, customers, shareholders, etc.</td>
</tr>
<tr>
<td></td>
<td>• Sex, age, etc.</td>
</tr>
<tr>
<td></td>
<td>• Role, position in organization, etc.</td>
</tr>
<tr>
<td></td>
<td>Which interpersonal relationships are of interest?</td>
</tr>
<tr>
<td></td>
<td>• Teams, group dynamics, structure of social networks, etc.</td>
</tr>
<tr>
<td></td>
<td>Which aspects of the institutional setting are of interest?</td>
</tr>
<tr>
<td></td>
<td>• Politics, power, technical system, complexity, interdependencies, etc.</td>
</tr>
<tr>
<td></td>
<td>Which aspects of the wider infrastructural system are of interest?</td>
</tr>
<tr>
<td></td>
<td>• Degree of competition, stability, etc.</td>
</tr>
<tr>
<td><strong>Interventions (I)</strong></td>
<td>What is the Intervention of interest?</td>
</tr>
<tr>
<td></td>
<td>• Leadership style, planning and control systems, training, performance management, etc.</td>
</tr>
<tr>
<td><strong>Mechanisms (M)</strong></td>
<td>What are the mechanisms of interest?</td>
</tr>
<tr>
<td></td>
<td>What is it about interventions acting in a context that leads to the outcome?</td>
</tr>
<tr>
<td></td>
<td>Why are mechanisms activated or not activated?</td>
</tr>
<tr>
<td><strong>Outcomes (O)</strong></td>
<td>What are the relevant outcomes?</td>
</tr>
<tr>
<td></td>
<td>What outcomes would be important to the individuals involved?</td>
</tr>
<tr>
<td></td>
<td>How will the outcomes be measured?</td>
</tr>
<tr>
<td></td>
<td>What is the primary outcome and what are the secondary outcomes?</td>
</tr>
<tr>
<td></td>
<td>• Performance improvement, cost reduction, low error rates, etc.</td>
</tr>
</tbody>
</table>

**Figure 39.1 Constructing review questions using the CIMO logic.**

In terms of internal dynamics and their experience of being a team member. It is, therefore, essential to define key terms during the question formulation phase. Since definitions and concepts are value laden, different perspectives and beliefs need to be made transparent. Although the review questions are specified a priori, a certain flexibility and modification of questions may be necessary as the reviewer gains a fuller understanding of the problem.

Until a very precise question(s) for a systematic review is specified, detailed work on a systematic review is best not started. It is impractical and unfair for novice researchers to begin a systematic review without a significant grounding in their field. Tranfield et al. (2003) suggest that prior to a systematic review, novices and less experienced researchers are best advised to produce a scoping study of the field. Scoping studies are exploratory reviews, since what to include and exclude is not self-evident. The purpose of the scoping study is to make clear the existing basis for the researcher's work and specify where and how the proposed systematic review fits into the current body of knowledge. The scoping study should help to define concepts and determine the review questions to be addressed.

**Step 2: Locating studies**

Systematic reviews aim to locate, select, and appraise as much as possible of the research relevant to the particular review question(s). The methods used to find studies (database searches, searches of specialist bibliographies, hand-searching of likely journals, and attempts to track down unpublished research) need to be reported in some detail.
Exhaustive literature searches of published and unpublished studies are conducted to ensure that the review results have taken into account all available evidence and are based on best quality contributions. The people of the search is to identify all relevant studies and can cover a range of different types of information (academic articles, books, practitioner materials, working papers, and the Internet). It sounds trite to say so, but the searches need to be aligned tightly with the review questions.

The search generally commences with an investigation of citation databases using search strings, grouping keywords, and applying search conventions.

- Simple operators include:
  - Truncation characters – ‘*’, ‘?’, e.g. team? – searches for documents which contain the term team, teams, teamwork, teamworking, etc.
  - Word association – ‘w’ or ‘near’; e.g. high(w)/reliability; high(near)reliability
  - Exact phrase – ‘’’, ‘’‘; e.g. ‘high reliability’; ‘high reliability’

- Boolean Logic includes:
  - OR, e.g. team* OR group* searches for either term in a document
  - AND, e.g. ‘high reliability’ AND organization searches for both terms in a document
  - AND NOT, e.g. team* AND NOT virtual* searches for documents which contain the term team, teams, teamwork, teamworking but not virtual

- By combining Boolean operators with parentheses complex searches can be constructed:
  - ‘high reliability’ AND (team* OR group*) AND NOT virtual

There is a common misconception that systematic reviews only involve electronic database searches. However, Greenhalgh and Peacock (2005) have shown that, in systematic reviews addressing complex questions, database searching can often account for as little as 30 percent of the total number of relevant articles. Therefore, a systematic search can be expected to use several methods, including: searching electronic databases, hand searching known journals, recommendations from experts, and cross-referencing. In addition to academic papers, articles in nonpeer reviewed journals, books, databases, websites, conference papers, seminars, workshops, technical reports and discussion papers and other ‘grey literature’ are all important. The decision whether or not to search specific data sources is dependent on the field and on the evidence available. For example, in nascent and fast moving fields, much of the evidence might reside in working papers. As with other aspects of the review, explicit and transparent choices must be made by the reviewer.

The output of the search is a comprehensive listing of articles and papers (core contributions) which helps to address the review questions. To aid information management, citation management software packages, such as Endnote or Refworks, can be used to store information accurately.

**Step 3: Study selection and evaluation**

Following the requirement for transparency of process, systematic reviews use a set of explicit selection criteria to assess the relevance of each study found to see if it actually does address the review question. Detailed decisions are recorded specifying precisely the basis on which information sources have been included and excluded. The aim of being explicit about the selection criteria is to make the reviewer’s decisions available for scrutiny and evaluation. Selection criteria also facilitate the updating of systematic reviews. One simple method for constructing selection criteria is to undertake a small number of some pilot searches in a citation database and then make a list of the
reasons for the inclusion and exclusion of articles. If the search is well focused then within a short time it is usually possible to produce an exhaustive list of reasons for inclusion and exclusion. The reasons for inclusion and exclusion then form the selection criteria.

General quality checklists, such as those produced by the Critical Appraisal Skills Programme (2008), may be used to help evaluate studies. Different subfields of management have different perspectives on research quality and so it is also important to take this into consideration. For example, some journals produce comprehensive guidance on appraising research evidence, such as the Journal of Occupational and Organizational Psychology’s guidance on qualitative studies. By combining general quality criteria and guidance from key journals in the field, it is possible to create a bespoke and practical quality appraisal tool. The purpose of the tool is not to exclude papers that are deemed to be of low quality, but to evaluate and report on the study and its limitations. However, if studies are excluded on the basis of quality, it is crucial for the reviewer to document and justify the reasons for this exclusion.

Step 4: Analysis and synthesis
After assembling the appropriate collection of relevant sources, the data analysis and synthesis commences. The aim of the analysis is to break down individual studies into constituent parts and describe how each relates to the other. On the other hand, the aim of synthesis is to make associations between the parts identified in individual studies. A synthesis needs to go beyond mere description by recasting the information into a new or different arrangement and developing knowledge that is not apparent from reading the individual studies in isolation.

The first step of analysis is to extract and store information on data extraction forms for every study included in the review. The data extraction forms generally address a series of interrelated questions, but the specific data collected may vary from study to study. The following questions were adapted from Wallace and Wray (2006) and Solesbury (2001):

1. What are the general details of the study – author, title, journal, date, language?
2. What are you seeking to understand or decide by reading this?
3. What type of study is this (philosophical/discursive/conceptual, literature review, survey, case study, evaluation, experiment/quasi experiment, etc.)?
4. What are the authors trying to achieve in writing this? What are the broad aims of the study? What are the study research questions and/or hypotheses?
5. How is the study informed by, or linked to, an existing body of empirical and/or theoretical research?
6. In which contexts (country, sector and setting, etc.) and which people (age, sex, ethnicity, occupation, role, etc.) or organizations was the study conducted?
7. What are the methodology, research design, sample, and methods of data collection and analysis?
8. What are the key findings?
9. How relevant is this to what we are seeking to understand or decide?
10. How reliable/convincing is it – how well-founded theoretically/empirically is this (regardless of method)?
11. How representative is this of the population/context that concerns us?
12. In conclusion, what use can I make of this?

The extraction form may also include the details of the reviewer (particularly if there are multiple extractors) and the date of the review. At the end of a systematic review, a full tabulation of all the included studies is displayed, providing a comprehensive summary representation of the field of study. By cross tabulating the studies, key issues can be identified. For example, in the column that lists the existing body of empirical and/or theoretical research that studies refer to, it will be possible to see whether the field has a single theoretical foundation or contending or complementary theories.

Many systematic reviews employ two or more independent reviewers to extract data from studies. When the interpretations and
findings of reviewers are compared, it is possible to minimize errors, resolve any differences, and produce a more robust data set. The aim of using data extraction forms is to provide an audit trail from the claims made in the review to the underlying evidence.

As noted in the foregoing section, reviews in management and organization studies are in general more likely to be interpretive or explanatory rather than aggregative. Harmonizing synthesis with the purpose of the review therefore is critical. In all cases, the resulting body of evidence is then explored, cross-tabulated and analyzed while engaging in rigorous reflection on any values, beliefs, and perspectives that might impact the interpretation. In a quantitative synthesis, the reviewer might highlight both the regularities and discrepancies in the data, whereas, in a qualitative synthesis, the reviewer could explore analogous and different meanings of respondents across the studies. Where multiple forms of evidence are available, it may be possible to attempt triangulation of the data/findings.

**Step 5: Reporting and using the results**

A systematic review is structured in a similar manner to a report of empirical research. The introduction section provides a statement of the problem and the review questions. The methodology section provides precise details of how the review was conducted – the search strategy, the selection criteria, and the analysis and synthesis criteria. The findings and discussion section contains a summary of all the studies in terms of the data extracted from the studies such as the percentage of studies in the field that are philosophical/discursive/conceptual, literature reviews, surveys, case studies, evaluations, or experiments/quasi experiments. The findings and discussion section also specifies precisely what is known and unknown about the questions addressed in the review. The conclusion section provides a summary of the review, the limitations of the study, recommendations for policy and practice, and future research needs.

In Cochrane-style reviews, review encouraged to think about whether or not there is sufficient evidence to provide clear lines for practice by asking three questions:

1. Will the practice improve outcomes?
2. Should the practice be abandoned in light of available evidence?
3. Are there trade-offs between known and known adverse effects?

If the review provides insufficient evidence to provide clear guidelines for practice, further questions are raised:

1. Is the practice promising but requires evaluation?
2. Does a practice that has been shown to have the effects expected from it require attention?
3. Is there reasonable evidence that practice is effective?

Reviewers in management and organization studies may make similar claims. However, since recommendations from research in our field are likely to be heuristic rather than algorithmic, it is essential that systematic review reports should provide considerable detail from the original studies so that users of the review can, if necessary, interpret the results and judge for themselves the strength of the findings.

**CONCLUSION**

Despite common misconceptions in the popular media, evidence-based practice has never sought to provide 'answers' to replace judgment and experience, instead attempts to inform decision making and action (Sackett et al., 1996). The Cochrane Handbook (Higgins and Green, 2006: 17) careful to point out that:

...the primary purpose of the review should be to present information, rather than offer advice. The discussions and conclusions should be to help readers understand the implications of the evidence in relation to practical decisions.
Academic evidence is only one input into the decision making process even in more developed fields such as medical science, for 'clinicians reason about individual patients on the basis of analogy, experience, heuristics, and theory, as well as evidence' (Cook et al., 1997: 380). Therefore, the terms 'evidence informed' or even 'evidence aware', rather than 'evidence based' may be more appropriate in management and organization studies (Tranfield et al., 2003). Hammersley (2004: 4) argues that:

...sometimes it will simply not be sensible to engage in elaborate explication of goals, to consider all possible alternatives, to engage in a search for information about the relative effectiveness of various strategies as against relying on judgements about this, or to try to measure outcomes. The linear rational model tends to underplay the extent to which in many circumstances the only option is trial and error, or even 'muddling through'.

However, any assumption that practitioners are unlikely to use evidence that has been produced by a systematic and rigorous process solely because they do not have the necessary time and appropriate skills, fails to question the desirability of such an outcome and draws attention to the training and development needs of managers. It is true that managers have at present rarely been trained to evaluate research but it is patronizing in the extreme to suggest that their efforts could not, or worst still should not be informed by the very best available evidence. For academic evidence to be used by managers it needs to be rendered accessible, palatable, relevant and useful. The linking of science base with practice remains a key challenge for management and organization studies. Therefore, in this chapter, we have identified alternative principles and outlined a methodology designed to further enhance and strengthen both management practice and management research. The methodology offers the potential to couple the best available academic evidence with the judgement and experience of practitioners in the true tradition of evidence-based practice.

REFERENCES


NHS Centre for Reviews and Dissemination. (2001) Undertaking systematic reviews of research on effectiveness: CRD’s guidance for those carrying out or commissioning reviews, NHS Centre for Reviews and Dissemination.


Shadish, W.R. and Sweeney, R.B. (1991) 'Mediators and moderators in meta-analysis: there’s a reason we don’t let dodo birds tell us which psychotherapies should have prizes', *Journal of consulting and clinical psychology*, 59(6): 883–93.


The Evidence for Policy and Practice Information and Coordinating Centre (EPPI-Centre), *Home* [Homepage of EPPI Centre], [Online]. Available: http://eppi.ioe.ac.uk/cms/ [February 2008].


