A Systematic Review of Comparative Studies of Attitudes to Social Policy

Trude Sundberg and Peter Taylor-Gooby
University of Kent, Kent, UK

Abstract

Systematic review (SR) is often promoted as a ‘best practice’ method to inform both policy-making and policy-evaluation in social policy in the light of the ever-growing volume of research. This article considers an innovative use of the method to advance and refine academic knowledge and illustrates this through a small-scale study of the literature on attitudes to welfare. SR relies on rapid, structured searches of large quantities of material. However, the method has encountered criticism. The article calls for a greater degree of reflection in terms of possible bias in SRs. A pilot using tools from SR methodology to survey attitudes towards social policy is used to demonstrate the problems. These include the US bias of major databases, and weaker reporting of book publications than of articles. SR may help to advance knowledge in social policy, but researchers need to be aware of its weaknesses and possible biases.

Keywords

Systematic review; Welfare; Attitudes; Social policy

Introduction

Systematic review (SR) has expanded rapidly during the past two decades as a means of evaluating research findings within topic areas. Advocates believe that SR has considerable value in summarizing key points across rapidly expanding fields of knowledge and is of particular relevance to policymakers. Critics see it as positivistic, limited and a creature of audit culture. The approach has been applied mainly to specific questions within relatively narrow fields of research in medicine, healthcare and education. However, the problems of dealing with large volumes of research activity and of understanding patterns, direction and the balance of evidence apply more generally in social science. The techniques and databases to enable rapid structured searches of large quantities of material have also developed rapidly. For these reasons, an examination of the potential for SR across fields where it has not

Author Emails: T.sundberg@kent.ac.uk; P.F.Taylor-Gooby@kent.ac.uk

© 2013 John Wiley & Sons Ltd
yet been applied is timely. We add to current knowledge by critically assessing
the use of SR as a tool for the evaluation of evidence, focusing on the first steps
of a review.

This article examines the value of SR techniques in investigating the
structure and directions of research across broader areas in social science to
advance academic knowledge. Problems may arise when SR relies too readily
on conveniently available sources of information. We take a middle position
between advocates and critics of the approach, and argue that SR is of value
in providing an overview of research and also in tracing issues in the way
knowledge develops within an academic field. SR may help synthesize existing
knowledge as a basis for further analysis. However, it must be applied caut-
iously so as not to induce a bias, taking into account the limitations of the
available material, the shortcomings of databases and the importance of books
versus articles in most social sciences. This article demonstrates the value and
limitations of SR in a review of cross-national comparative studies of attitudes
to social policy. It falls into three sections: the first is a discussion of recent
debates about SR; the second is an analysis of the strengths and weaknesses of
the approach, illustrating how these strengths and weaknesses emerge in
relation to social science through application of the method in a pioneering
SR review of cross-national attitude surveys; and the third is a demonstration
of the value of SR when used with caution in a presentation of substantive
findings from the review. We draw conclusions about the field and about the
potential for using SR more broadly in substantive areas of social science as a
technique for the evaluation of evidence.

**Systematic Review**

SR could be described as ‘one of the success stories of the 1990s. In the space
of ten years the movement has had a significant impact on health care and
policy’ (Trinder and Reynolds 2000: 1). Some see its growth as a result of the
development of meta-analysis as an alternative to traditional, narrative reviews
of research literature (Hammersley 2001: 543). Others point to the increasing
pressure to ensure that policy and practice rest securely on research evidence
(Davies et al. 2000). Policymakers themselves increasingly seek research
projects that can provide input and advice to help shape new policy interven-
tions. SR has become a minor academic industry flourishing in the climate of
evidence-based policy-making and driven forward by the demands for more
rational resource allocation in healthcare, education and elsewhere, epito-
mized by the work of the National Institute for Clinical Excellence. It formed
the core of the invitation of David Blunkett (Secretary of State for Education
and Employment) to ‘social scientists to work with us to find out what works
and why and what types of policy initiatives are likely to be most effective. This
is crucial to our agenda for modernising government’ (Blunkett 2000). Aca-
demic engagement and research funding have established an institutional
framework. The Economic and Social Research Council (ESRC) promotes
work on SR through the National Centre for Research Methods (NCRM
2010). The UK Cochrane Centre and the EU- and ESRC-funded Evidence for
Policy and Practice Information and Co-ordinating Centre were both set up in
1993, the National Health Service Centre for Reviewers and Dissemination (CRD) was set up in 1994 and the Campbell Collaboration in 1999.

Formal definitions of SR highlight structure, focus and transparency, ‘reviews of this type: use a formal protocol to guide the process; focus on a specific question; identify as much of the relevant literature as possible; appraise the quality of the studies in the review and synthesise their findings’ (Grayson and Gomersall 2003: 3). They rest on the assumption that it is possible to achieve unambiguous, comprehensive and unanimous agreement on the quality of a large number of research projects and apply this to achieve results:

A systematic review has clearly identified questions, searches for relevant research following a procedure specified in advance, has criteria for which studies to include, has criteria for the information to be considered from these studies, appraises studies using clear criteria for what is good and what is less good research and synthesises the results in a transparent fashion, sometimes using statistical pooling (Bradshaw et al. 2000: 6).

From one perspective this is simply the application of reason to the evaluation of research evidence, ‘systematic reviews are needed because they accumulate evidence from multiple studies which may prevent practitioners and policymakers from basing their decisions on single studies that might be flawed and biased’ (CRD 2009: v). An alternative approach analyses SR from a sociological viewpoint as a shift in intellectual practice which contributes to the restructuring of the way research evidence is understood and thus to the way in which it can influence policy-making. In this view, one of the most striking developments during the past 30 years has been the development of a culture of audit and risk avoidance (Power 2004). As Pawson and Tilley (1997: xii) argue, ‘we live in a knowledge-centred, value-adding, information-processing, management-fixated world which has an obsession with decision-making’. Analysts of risk society argue that individuals are now both better-informed and less deferential to experts (Beck 1992: Chapter 1; Wynne 1996). As a response to these societal changes, SR provides a technology to maximize the value generated from research and, at the same time, to demonstrate the authority of the evidence that informs policy. The point of SR is summed up in the claim that, in a less respectful world, ‘organised distrust generates trustworthy reports’ (Campbell 1984: 38).

Critics of SR often emphasize four issues: the principle of replication, the role of judgement, the use of a hierarchy of evidence, and the failure to understand the contextual nature of knowledge.

Replication is one of the defining principles of a systematic review, and is achieved through rigorous documentation and by making every step of the review transparent. One of the most influential critics of SR, Ray Pawson, argues that it is impossible to account for and document every point of a review. Furthermore, procedural uniformity may be counterproductive as flexibility and iteration may be crucial to appreciate the variety of evidence found. He advances a ‘model of validity that rests on refutation rather than replication’ (Pawson 2006: 26), as does this article.
Due to the wide range of methodological approaches and less standardized research processes in social science compared to natural sciences, the replication principle combined with hierarchical evidence approaches may exclude and simplify rather than systematize and enhance knowledge. We underline the importance of judgement in including and interpreting heterogeneity in studies, instead of eliminating it as an anomaly in the way that some meta-analysis software does (Petticrew and Roberts 2006: 197). This is because heterogeneity in itself may be valuable and expected in social policy. SR in this area needs to recognize this or risk reinforcing a skewed research synthesis which does not take real-world complexities into account.

Many studies organize findings according to a hierarchy of evidence, determined by claims about internal validity. From this perspective, SR and meta-analysis are given the greatest weight, followed by randomized controlled trials (RCT) with definitive results and then by RCTs which fail to generate definitive conclusions (Petticrew and Roberts 2006: 58). Knowledge obtained by other means (e.g. quantitative survey or qualitative analysis) is accorded correspondingly less weight.

This approach follows a natural science model and builds on techniques developed to test the value of interventions in medical settings where everything beyond a limited range of measured variables is assumed to be either controlled or irrelevant. Critics suggest that important research may be ignored (Hammersley 2001: 546; Evans and Benefield 2001) and the diversity of methods reduced as research is driven towards the medical model (Evans and Benefield 2001). They underline the importance of diversity in social sciences’ methods that may bring about heterogeneity in a review, and argue that this should not be seen as a weakness but as contributing to understanding. More recently, manuals on SR have encouraged practitioners to consider explicitly ‘which type of study is most appropriate for answering your review question’, and have encouraged reviewers to take account of studies using a range of methods, including qualitative approaches, as appropriate (Petticrew and Roberts 2003: 59). However, positivistic models of replication and transparency still tend to predominate.

The fourth criticism underlines the need to consider context when assessing evidence of the effectiveness of an intervention (Kitson et al. 1998; Pawson 2006). SR traditionally focuses on ‘what works’ without taking into account the context in which interventions have taken place. Realist synthesis seeks to include these factors. It is concerned with explanation and is driven by the question ‘What works for whom in what circumstances and in what respect?’ (Pawson 2006: 18). It rejects the principle of replication and emphasizes that of refutation: studies are considered not so much in terms of whether they reinforce a consensus as whether their findings are disproved. We argue that issues of context also apply to the academic and scholarly context in which studies are shaped and conducted, for example in terms of dominant schools of thoughts and methodological approaches, which may shape their findings. Pawson argues that research studies are conducted in a complex social context and are never carried out in exactly the same way or in exactly the same setting (Pawson 2006: 6).
SR essentially sums the number of studies in the literature which point to specific conclusions and which meet particular quality criteria, often resting on assumptions about the hierarchy of evidence generated through particular methods. Realist synthesis argues, instead, that this process may discard evidence that is of value, despite limitations of method, and which can be used to support diverse interventions and programmes as appropriate in various contexts (Pawson 2006: 12). It sees SR as an iterative process that should be driven by scepticism and the principle of refutation, and where interpretation and theory is crucial. We agree with Pawson’s point that the judgement of the researcher cannot be written out of the process, but must be made explicit and transparent:

Science is more than excellence of execution, more than durability of data. Neither is it simply a product of following a pre-determined investigatory pathway or its course would be mapped out already. The stark, staring implication for this study is that there will never be a simple litmus-test for science; there are no instantaneous warrants for declaring certain procedures as valid science (Pawson 2003: 6).

SR has generated enormous interest at the interface between social science and policy-making. The above comments indicate that there are strong grounds for considering the appropriateness of particular methodologies and kinds of evidence in particular contexts, and that there is controversy about the limitations to positivism. SR needs to ensure that it does not exclude studies that point to relevant findings, while retaining a critical perspective that enables it to eliminate methodologically unsatisfactory work and appreciating the heterogeneity found in social science. It remains helpful for researchers to take an overview of work across a field, discover the key patterns within it, point to areas where work is concentrated and where there are fewer studies and report the chief findings within it, point to areas where work is concentrated and where there are fewer studies and report the chief findings supported by various approaches. As the Social Care Institute for Excellence argues, ‘Systematic review methods can be applied to any type of question . . . transparent and replicable methodology should be applied to all forms of literature review in the interests of quality and reliability’ (Coren and Fisher 2006: 2).

In this article we seek to apply techniques from SR in a broader field, the comparative analysis of attitude surveys. We use this work to illustrate problems with the use of SR in social policy. We focus on the early stages of SR – data collection rather than interpretation – since these processes have tended to receive less attention. The article ascribes to a position where explanatory schemes are not ‘arrived at on the basis of the data, but rather on the basis of knowledge, theoretical formulations and assumptions, and logical analysis’ (Pedhauzer 1983: 579). SR cannot substitute theory.

We argue in favour of the importance of an iterative process and of judgement, context and scepticism when conducting SR in social policy to acknowledge the diversity of the field. We manage large volumes of data through the use of random sampling. We reject assumptions about a hierarchy of methodology and seek to include the widest range of research, evaluated as appropriate to the various studies. However, the strengths and
weakness of particular pieces of work must be taken into account. We also argue that the assumptions about replication and about formalized protocols often made in SR reviews of more technical literatures may be problematic in a social science context. Such issues may arise in more narrowly focused and natural science dominated fields due to differences in the ways databases are constructed and particularly in the rules governing the presentation of information in journal abstracts. They are much more relevant in social science due to differences in vocabularies and the definitions of terms between writers, the lack of any common structure to abstracts and the gaps in coverage of databases. These problems require a broad and pragmatic approach as our empirical work shows.

SR is not a royal road to knowledge, it is simply a convenient technique for the review of large numbers of studies, ‘that aims to provide an account of the literature that is comprehensive, capable of replication and transparent in its approach’ (Bryman 2008: 700). This approach does not follow assumptions about the hierarchy of methods. It requires considerable amounts of time to be spent in evaluating the contribution of particular studies. Its value is that it points to methods for systematically searching databases in order to reduce bias in the selection and management of findings from a number of studies.

This leads to an important pragmatic concern. Our research indicates that there are real shortcomings in the most conveniently available databases – some important material is omitted, thereby damaging literature reviews based on them. This suggests that SR may generate problems that are not immediately obvious to those using the approach less critically. These problems need to be addressed by using as wide a range of databases as possible and by supplementing them from other sources, rather than simply applying a protocol mechanistically to the material available from the most convenient sources. Transparency about the techniques used to generate the review is essential for the outcome of the approach to be fully understood by the academic community.

Illustration of Strengths and Weaknesses: The Attitude Survey Study

In this section we discuss attitude survey-based studies which serve as an illustration of potential problems with SR. This is a suitable area for examination because it is highly international and spans several databases. It illustrates the issues surrounding the adequate recognition of material from a range of sources. It has also expanded rapidly as an area of study in recent years. The studies for the most part centre on a relatively small number of major international surveys, so that there is a strong need for systematic comparison of different interpretations of the data. SR can contribute to this. We recognize that a pilot study may have some methodological shortcomings due to limitations of scope and sample (Teijingen and Hundley 2001), but show that the sample is an adequate foundation for a critical discussion of some central issues in applying SR to social policy.

Our aim in this research was to cover the full range of material generated in the field, ensuring that the studies included are comparative and cross-
national and deal with material from attitude surveys. For practical reasons the research was limited to studies published in the English language covering Europe and the USA, Canada, New Zealand and Australia where the study also included European countries. Expert advice suggests that the most important omission is the German-language literature, but that the vast majority of published work is included.

To cover the full range of publications, the review included the literature listed in the International Bibliography of the Social Sciences (IBSS) and ISI Web of Science (WoS) databases (which are largely limited to articles in the journals included by the databases, with some conference papers), plus hand searches of the reference lists from the chapters of a recently published book which reviews the field and a research paper to control for differences in items identified through this method. We chose these two databases to include the leading international database with coverage across a wide range of disciplines, and a database more specifically focused on social science. We did not include Elsevier’s Scopus database due to its natural science focus, ‘two per cent of the material in Scopus is from social science, as against 14 per cent in WoS’ (Jasco 2005: 1540). For this reason, and since we were under time pressure, we chose one database with wide coverage of subjects and journals and one social science-focused database: WoS and IBSS. As Jasco argues, ‘WoS comes the closest to be a genuinely pan-disciplinary database’ (Jasco 2005: 1540). According to his calculations the distribution of records among the major component of WoS is: science 77 per cent, social sciences 14 per cent, and arts and humanities 9 per cent; whilst for scopus it is: health and life science 60 per cent, chemistry, physics, maths and engineering 25 per cent, biological, agricultural, earth and environmental sciences 13 per cent, and social sciences, psychology and economics 2 per cent (Jasco 2005: 1540).

Ninety-three search terms were identified in the field of welfare, state provision, welfare finance, citizenship, poverty, inequality, need, social justice and related normative issues, and combined using Boolean operators (see Appendix 2). The list of terms was refined from a longer list through an iterative process of consultation with 15 academics in social policy, sociology, criminology, economics, politics and psychology at the University of Kent. The search was carried out in November and December 2009 covering the 15-year period between 1994 and 2009 and achieving more than 12,000 hits. Figure 1 summarizes the SR process. The material from each database and that identified through the lists of references from books and chapters in the field was screened twice, once using the keywords and abstracts provided by the database and once against the inclusion criteria using the full study and the judgement of the researcher. Duplicate material was identified and removed. This follows the procedure recommended by the CRD (2009: 13).

As can be seen from figure 1, the initial search generated a high proportion of non-relevant hits, nearly 99 per cent being excluded by the screening process. One reason for this was the use of an inclusive search protocol that ranged over a considerable area in social science. Examination of the Cochrane and Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) databases indicates that proportions as high as this are common in more tightly-focused reviews in the medical and
educational literature. In any case, a low proportion of relevant studies is to be expected in social science due to the fact that many journals do not impose rigour in titles, do not use keywords and do not insist on the inclusion of methodological information in abstracts (Petticrew and Roberts 2006).

**Key findings**

Five points emerge from the review and in studies of the databases:

- databases tend to over-represent US and under-represent other literatures;
- there are substantial differences in coverage between the different databases;
- WoS, with its assumptions and biases, tends to dominate database work;
- the vocabularies of key search terms tend to differ between disciplines so that a search from one perspective may miss relevant work from another perspective; and
books, book chapters and research papers are systematically under-represented in all databases, despite being of considerable importance in many areas of social science.

We will present an overview of these issues and then illustrate and refine them, using the material from our own systematic review of the literature on attitudes to welfare. Our findings emphasize practical issues underlining the importance of using an iterative research process in which knowledge and judgement are crucial at every step when conducting SR in social sciences.

**US over-representation**

First and perhaps most importantly, much of the data available is biased towards the USA. This reflects the use made of the WoS database, which has much better coverage of US journals. This points the importance of gate-keeping processes regulating the inclusion or exclusion of journals from the databases. Braun and Dióspatony (2005), analyzing natural science publications, show that the USA, which has 54 per cent of the gate-keepers, contributes 32 per cent of the papers included. The UK is in second place with 10 per cent of the gate-keepers and correspondingly fewer papers, only 8 per cent (Braun and Dióspatony 2005: 1549). On the other hand Almeida et al. (2009) argue that size of country and size of institution influence productivity, in terms of published papers, and thus numbers of citations, ‘this is consistent with recent observations from several quality and efficiency criteria showing that the United States are outperforming Europe in terms of science’ (Almeida et al. 2009: 135). About 36 per cent (4,142) of the titles found on WoS are US titles. Seventeen per cent (504) of IBSS titles are from the USA (numbers from JISC, Academic Database Assessment Tool) again showing the difference in coverage of the USA, in particular for the biggest database, WoS. Scholars need to take this bias into account in reading and conducting SRs.

**Differences between databases**

The databases differ substantially in their coverage of titles from different European countries (see table 1). This is of interest since comparative attitude studies are published in a number of countries and it is necessary to consider whether there are different national traditions. IBSS is rather weaker than WoS in German journals but particularly strong in those from France. Another interesting finding is the low overlap between WoS and IBSS in the titles covered: 16 German titles out of 368 are identified in both databases and 32 out of 839 French ones. Limiting inclusion to studies covering European countries ensured comparability between the social science surveys which the studies analyze. However, this creates problems when there are different national traditions in social science. Databases that vary in their coverage of a country’s scientific contribution may bias the reviews based on them. For example, a comparison of the top 20 Polish sociologists cited in the Social Sciences Citation Index (SSCI) and a Polish index had only 12 names in common (Hicks 1999: 203). Hicks argues that this is due not only to the fact
that some journals are not included in international databases, but also to the fact that studies with a national character, not addressed to an international audience, may get less attention. However, one might argue that any analysis of a phenomenon might be interesting and of use to others studying a similar theme or applying similar methodologies.

National differences in terms of coverage, and in the production of articles on certain topics, may result in a bias in the material included both in systematic reviews and other literature reviews. Citations and databases may not reflect the most important contributions and authors within a field, as Hicks shows. Citation rankings may influence the financing of institutions, judgements of the quality of a university, and evaluations of the strength of research in a country (Braun and Dióspatonyi 2005: 1548).

The databases searched will have an impact on what evidence is found in an SR. This underlines the importance of judgement and context when interpreting findings, and of an iterative process where changes may be made to a review protocol during the research process.

**The dominance of Web of Science**

The third point is the domination of WoS in the process of SR and in its coverage of a vast amount of journal titles. WoS provided the majority (8,332) of all initial hits in the attitude survey, roughly twice as many (4,048) as from IBSS, as can be seen in table 2. There was substantial duplication between the databases: 141 out of 236 items were excluded as duplicates when they were combined. The high quantity of hits made when searching WoS reflect the large number of journal titles covered by WoS so that it is essential to include this database in any review.

However, exclusive reliance on WoS may be misleading. To analyze the dominance of WoS, the JISC Academic Database Assessment Tool, which facilitates comparisons of databases, was used. The testing was done on WoS rather than Web of Knowledge (WoK), as WoS also includes conference proceedings, patents and open access material (http://wokinfo.com/about/faq/ [accessed 13 June 2010]) but the tool gives a good insight in the different coverage of the two databases. Table 3 shows that the two databases vary

<table>
<thead>
<tr>
<th>Database</th>
<th>English</th>
<th>German</th>
<th>French</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBSS</td>
<td>65%</td>
<td>7%</td>
<td>28%</td>
</tr>
<tr>
<td></td>
<td>1,570</td>
<td>161</td>
<td>680</td>
</tr>
<tr>
<td>WoS</td>
<td>96%</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>9,398</td>
<td>207</td>
<td>139</td>
</tr>
</tbody>
</table>

*Source: JISC, Academic Database Assessment Tool, data from 2010.*
dramatically. WoS is roughly four times as large as IBSS, including 11,456 titles against IBSS’s 2,712. The limited overlap is also striking. Less than 10 per cent of WoS titles are included in IBSS, making up 40.7 per cent of that database’s titles. IBSS is clearly more centred on European (including UK) literature (table 3). This reinforces the earlier point that researchers should search more than one database to guard against bias. It is also important to include books for acceptable coverage of the areas in question. WoS does not include all the titles in IBSS, again underlining the importance of going beyond this database.

Table 3 also shows that IBSS has better coverage of Europe than WoS, but a more limited coverage of US material. Searches limited to WoS risk finding out what works only in a US setting.

**Different traditions and search terms**

It is advisable to include wide-ranging terms in order to ensure that the topics are fully covered. One issue that emerged was the relevance of terms that have
emerged in other disciplines but are at the periphery of social policy debates (e.g. ‘social capital’, ‘citizenship’, ‘legitimacy’, ‘identity’) to the current interests of the subject (see Grayson and Gomersall 2003 for a discussion of changing emphases in social science literatures). The field is relatively broad, with changing emphases, and accommodating academic interests from different directions.

The under-representation of books

Substantial differences emerged between material identified from the screening of books, chapters and papers, and that from the databases which included almost exclusively articles (with some conference papers covered by WoS from 2009). This shows the importance of including such material in database searches in social science (see Hicks 1999: Table 4 ‘Overlap between IBSS and WoS’). ‘The centrality of books in scholarly communication in the social sciences contrasts with their absence in literature databases, including the SSCI . . . journal based social science indicators will be problematic given the heterogeneous literature of the social sciences’ (Hicks 1999: 201–2).

This point is rarely discussed in SR literature but matters if the technique is to be extended outside narrow medical and science-based literatures. Almost 10 per cent of the references culled from book and chapter reference lists were included in the final study after the two waves of screening, as against less than 1 per cent of those from the databases. The 112 book references remaining after the second screening included only 41 duplicates (37 per cent), compared with 141 out of 236 (nearly 60 per cent) for the databases. When the final material from the book references (71 items) and the databases (95 items) were combined, only 14 duplicates were found, less than 10 per cent.

The different databases cover substantially different literatures. WoS is much larger and US focused, IBSS has more on Europe and particularly on French literature. In addition, there is much material in reference lists, especially in books, that must be included because it is not covered in the databases. Thus an unbiased SR needs to use the full range of available databases, to be alert to overlaps and also to include as much material as possible from non-database sources, such as reference lists from book chapters.

We now develop these points further in the context of a small-scale study of attitudes to welfare, carried out to illustrate the potential of SR for bringing together material across a broader field to develop knowledge in social science.

A Demonstration of the Value of Systematic Review in Social Science and of the Issues of Quality and Bias

Our study was carried out to illustrate and develop methods. It emphasizes the importance of judgement and interpretation and, therefore, the laborious and time-consuming nature of the review exercise. For this reason, further analysis was based on the detailed review of a sub-sample of 24 studies, chosen at
random from the studies remaining after the second stage of exclusion. A preliminary review of the content of the sample considered methodology in order to eliminate material where the method was unable to support the findings. However, some studies give few details of method, especially of justification for their choice of countries, and do not discuss issues of sample size and populations within countries. Second, technical advances, particularly the spread of multi-level modelling, time-series and fuzzy logic analyses, enable researchers to interrogate the data more searchingly. Since there is no commitment to a methodological hierarchy of evidence, studies were considered on their merits, in relation to the questions they sought to answer. Third, most researchers were limited to the analysis of the standard questions provided by major cross-national surveys such as the International Social Survey Programme, Euro-Barometer, the World Values Survey, the European Values Survey and the European Social Survey (see Appendix table A1). These questions were often interpreted rather differently and applied to different themes and hypotheses. Again, commentary is based on the merits of the individual study.

The review of the content of the papers identified a number of points which reinforce and extend the general findings discussed in the previous section. These points concern the scope and orientation of studies and the importance of disciplinary perspective in this process.

Explanatory factors: individual characteristics and regime type

Studies often treat attitudes as related to individual characteristics (social class, gender, age, faith, political orientation, and so on) and features of welfare states (regime type, constitutional and institutional structure, variety of capitalism, security of employment, gender policies, and so on). In general, the studies treat individual characteristics as embedded within the macro-level features of welfare states. Multi-level modelling provides a methodology which enables researchers to take this approach much further and which is now being enthusiastically embraced (Van Oorschot and Uunk 2007; Esser 2005; Van Groezen et al. 2009; Anderson and Singer 2008). The most important macro-level characteristics used in the research concern regime type, following the model associated with Esping-Andersen (1990). Seven out of the 24 studies were based explicitly on tests of this approach and all found some confirmation of it.

The macro-level features of welfare states were used to combine characteristics in different ways and to examine different issues. Alesina and Glaeser (2003) found that ethnic heterogeneity has the strongest explanatory power in relation to level of social spending, while political institutions, geography, and presence of war also influence redistribution. Vala et al. (2004) found that whether a country had a tradition of emigration or immigration had an impact on solidarity with immigrants, while Van Oorschot and Uunk (2007) found that the proportion of immigrants in a country affects the level of public concern about immigration. Anderson and Singer (2008) show that level of inequality in a country impacts on satisfaction with the political system. Iversen and Soskice (2001) found that income equality, social spending and a
nation’s education system are related all help shape skills sets which, in turn, influence level of social policy preferences. Esser (2005) tested a range of macro-level indicators and found that regime type and unemployment rates influence attitudes to retirement.

Huseby (1995) showed that economic differences between countries and gross domestic product per capita both have explanatory power when analyzing attitudes about size and responsibility of governments. Roller (1995b) found that countries, standards of living, income inequality, wealth of a country and political agendas all have effects on attitudes towards socio-economic equality. Newton and Confalonieri (1995) and Pettersen (1995) found that political circumstances and national political mood impacts on support for paying taxes and attitudes to socio-economic security.

The influence of academic disciplines

Analysis of the studies by discipline of researcher shows a general pattern. Those from political science tend to focus on the political sphere and political circumstances rather than the welfare state system and welfare outcomes (e.g. Borre and Scarborough 1995; Iversen and Soskice 2001; Anderson and Singer 2008; Gijsberts and Nieuwbeerta 2000). Economists tend to analyze attitudes to welfare in relation to self-interest and the provision and finance of benefits (Alesina and Glaeser 2003; Van Groezen et al. 2009; Iversen and Soskice 2001), while sociologists pay more attention to social contacts, social exclusion and deservingness (Mau 2003; Van Oorschot and Uunk 2007; Roller 1995a, 1995b).

The review of the content of the articles underlines the importance of a theory-led approach to synthesis to incorporate the heterogeneity of the studies and to use different findings from different approaches to advance knowledge. In a larger study, currently in train, this approach will be further developed in order to map the range and structure of attitudes, the circumstances in which they are found and who displays them.

Conclusions

Taken together, the findings from the general review of the databases and the study of the attitudinal literature reinforce the point that different researchers pursue cross-national comparative attitudinal studies from different perspectives and in order to answer different questions. Most SR is directed at identifying what social science research has to say about a specific problem or issue. It seeks the answer or answers that are best supported by the balance of evidence. The domain we have examined is broader and less conveniently structured. The research supports two points relevant to the extended use of SR techniques as methods for evaluating evidence in order to advance academic literature and theory-building rather than as simply a tool for assessing ‘what works’ in a defined context as a guide to policy-making.
• First, the most widely used databases are limited to particular literatures and differ in coverage. Searching within them suffers limitations in their capacity to identify material relevant to European contexts. Academic work must approach databases critically, ensure that a range of databases are included and also consider literature outside the databases as currently constructed.

• Second, books, book chapters and other material offer important sources of information that is, to some extent, independent from that contained in the articles that comprise by far the greatest proportion of the databases. SR review may be misleading if it focuses exclusively on the material conveniently available in electronic databases. There appears to be no current substitute for hand searches of bibliographies. It is possible to combine these with SR of the databases, provided these are approached critically and the hand searches are conducted systematically according to similar protocols.

• There is a risk that uncritical use of the electronic databases in SR academic work (and more broadly in literature reviews) may produce a perspective limited to material published in journals, which does not cover all of social science knowledge and research, although it includes an important part of it, and fails to give adequate weight to any differences between US and other literatures.

Further points concern the substantive study of attitudes to social policy. Discipline pays a noteworthy role in directing the interests and approach of researchers. There is considerable variation in practice across the field. SR shows that comparative cross-national attitude studies, despite the number and sophistication of relevant publications, have yet to develop into a domain of study where it is possible to point to shared findings that provide a secure foundation for hypothesis-building and further development. The ideas tested come from other areas and reflect wide-ranging interests.

SR is likely to play an increasing role in evaluating social science evidence as the quantity and range of material to be synthesized grows, as it comes to be structured and cross-indexed within databases, and as databases enlarge their coverage. It has the potential to be extended from closely policy-related work to more theoretical and academic studies where the capacity to synthesize a very large range of material is attractive. We have shown a number of weaknesses in naïve approaches to SR particularly concerning databases: that different databases over-represent and under-represent different literatures; that all tend to ignore one of the more important channels of social science communication – books; and that disciplinary perspective shapes both the orientation of studies and of searches. We hope that these points will serve both as words of caution and also as encouragement to future researchers to use these methods alongside others, and to employ them critically, as they would other methods.
Appendix 1

Table A1

Cross-national surveys used by the studies reviewed

<table>
<thead>
<tr>
<th>Survey</th>
<th>Number of studies using the survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESS</td>
<td>2</td>
</tr>
<tr>
<td>ISSP</td>
<td>10</td>
</tr>
<tr>
<td>Eurobarometer</td>
<td>9</td>
</tr>
<tr>
<td>ISJP</td>
<td>1</td>
</tr>
<tr>
<td>ISEA</td>
<td>1</td>
</tr>
<tr>
<td>Political Action</td>
<td>3</td>
</tr>
<tr>
<td>EVS</td>
<td>1</td>
</tr>
<tr>
<td>WVS</td>
<td>2</td>
</tr>
<tr>
<td>Other cross-national</td>
<td>2</td>
</tr>
<tr>
<td>National surveys</td>
<td>6</td>
</tr>
</tbody>
</table>

Appendix 2: Search Strategy

(Attitudes AND (Welfare OR Religion OR Trust OR Political efficacy OR quality of service OR public officials OR political authorities OR political actors OR political system OR Political participation OR Social capital OR Individuals beliefs OR personal beliefs OR Postmaterial beliefs OR Citizenship OR Civic duty OR Civic virtue OR Civil liberties OR Deviance OR cheating on taxes OR bribery OR Abuse of welfare OR Social benefits make people lazy OR Pension OR Pensioners OR care for the elderly OR Ageing OR Class OR Poverty OR Social exclusion OR Entitlement to benefit OR State responsibility OR individual responsibility OR Government intervention OR ownership of social service institutions OR welfare support OR more government spend OR Less government regulation OR tax OR progressive taxation OR government spend OR welfare regime’s generosity OR cuts and reductions in public services OR service delivery by private companies/civil society OR welfare provision by the state OR financing welfare OR Responsibilities of EU OR Employment OR Social protection OR Social justice OR Distributive justice OR Redistribution OR Inequality OR Opportunity OR Fairness OR Solidarity OR Social legitimacy OR legitimacy OR Social inclusion OR Social cohesion OR Risk OR Religious identity OR social identity OR Class identity OR National identity OR European identity OR Family OR new forms of family life OR traditional forms of family life OR Maternity leave OR Parental leave OR Duty of women OR Children OR Youth welfare OR Gender OR Women’s rights OR Mental health OR Immigration OR migration OR Diversity OR Intolerance OR Cultural integration OR integration OR race OR disabled OR long-term sick benefits OR Housing benefits OR Future of welfare OR Future of public sector OR Financial concerns OR Confidence in the economy) AND Survey NOT (Japan OR Russia OR China).
Searches were limited to English language publications, published 1/1994–12/2009.

References


Centre for Reviews and Dissemination (CRD) (2009), Systematic Reviews: CRD’s guidance for undertaking reviews in healthcare CRD, York: University of York.


Hicks, D. (1999), The difficulty of achieving full coverage of international social science literature and the bibliometric consequences, Scientometrics, 44, 2: 193–215.


**Systematic review institutions and resources**


