Systematic review methodology in higher education

Margaret Bearman, Calvin D. Smith, Angela Carbone, Susan Slade, Chi Baik, Marnie Hughes-Warrington & David L. Neumann

To cite this article: Margaret Bearman, Calvin D. Smith, Angela Carbone, Susan Slade, Chi Baik, Marnie Hughes-Warrington & David L. Neumann (2012) Systematic review methodology in higher education, Higher Education Research & Development, 31:5, 625-640, DOI: 10.1080/07294360.2012.702735

To link to this article: https://doi.org/10.1080/07294360.2012.702735

Published online: 09 Oct 2012.

Submit your article to this journal

Article views: 3162

Citing articles: 18 View citing articles
Systematic review methodology in higher education

Margaret Bearmana*, Calvin D. Smithb, Angela Carbonec, Susan Sladec, Chi Baikd, Marnie Hughes-Warringtone and David L. Neumannf

aHealthPEER, Monash University, Melbourne, Australia; bGriffith Institute for Higher Education, Griffith University, Gold Coast, Australia; cOffice of the Pro-Vice Chancellor (Learning and Teaching), Monash University, Melbourne, Australia; dCentre for the Study of Higher Education, University of Melbourne, Melbourne, Australia; eAustralian National University, Canberra, Australia; fSchool of Applied Psychology, Griffith University, Brisbane, Australia

Systematic review methodology can be distinguished from narrative reviews of the literature through its emphasis on transparent, structured and comprehensive approaches to searching the literature and its requirement for formal synthesis of research findings. There appears to be relatively little use of the systematic review methodology within the higher education sector. This paper outlines the systematic review methodology, including variations, explores debates regarding systematic reviews from the educational literature and describes particular issues for its application within higher education. We conclude that thoughtful use of the systematic review methodology may be of benefit to the sector.

**Keywords:** higher education; research methodologies; research synthesis methodologies; systematic review

**Introduction**

Systematic review methodology is a protocol-driven and quality-focused approach to summarising healthcare evidence. Over two decades old, this methodology has revolutionised healthcare delivery, funding and research and is strongly associated with the now ubiquitous phrase ‘evidence-based practice’. The methodological standards are supported by the work of an international not-for-profit organisation called the Cochrane Collaboration, which ‘prepares, maintains and promotes systematic reviews to inform healthcare decisions’ (Cochrane Collaboration, n.d.).

Evans and Benefield (2001) called for the use of systematic reviews within education practice and policy, illustrating the debate with a description of their own systematic review on interventions to support primary students with emotional and behavioural difficulties. In reflecting on their experience, they describe their work as the equivalent of any significant research project and conclude that systematic review is an excellent mechanism for revealing the gaps in current research and providing direction for future funding decisions.

Since the late-1990s, the systematic review has been the literature review methodology of choice for certain sectors of the educational research community, led by organisations such as the Campbell Collaboration, The Evidence for Policy and Practice...
Information and Co-ordinating Centre (EPPI-Centre) and Best Evidence Medical Education (BEME). Medical, nursing and other health educational researchers, in particular, consider systematic review as a key methodology. For example, in 1999 the prestigious *Journal of the American Medical Association (JAMA)* published a systematic review on the effects of continuing medical education methodologies on doctor behaviours or patient outcomes (Davis et al., 1999). A broad indicator of the regard for the *JAMA* systematic review is shown by the 849 publications citing this paper, as indicated by a 2011 citation search on the Web of Knowledge. Examples of other more recent systematic review topics in the health professions include high fidelity simulation (Issenberg, Mcgaghie, Petrusa, Lee Gordon, & Scalese, 2005), reflective practice (Mann, Gordon, & MacLeod, 2009) and workplace assessment (Miller & Archer, 2010). In other parts of the educational research community there has been a less warm response. In particular, there is a long-running debate about the relative value of a methodology that has its foundations in the positivist paradigm, especially that built upon quantitative measurement of efficacy (Clegg, 2005; Denzin, 2009; Hammersley, 2001).

An examination of the higher education literature indicates the paucity of systematic reviews published within higher education journals outside of the health professional education domain. An Educational Research Information Clearinghouse (ERIC) search for the term ‘systematic review’ or ‘systematic literature review’, coupled with synonyms for ‘higher education’, yielded 16 peer-reviewed journal articles, of which five were systematic reviews in higher education (Gough, Kiwan, Sutcliffe, Simpson, & Houghton, 2003; Jackson & Kile, 2004; McGrail, Rickard, & Jones, 2006; Spelt, Biemans, Tobi, Luning, & Mulder, 2009; Trede, Macklin, & Bridges, 2012). The limited use of systematic review within the higher education field is an interesting phenomenon as it is a methodology well used in other educational research sectors. As Evans and Benefield (2001) note, systematic reviews are time-consuming, but this does not seem in itself a sufficient hurdle.

This paper critically reflects on the recent and potential use of systematic review in higher education research, a discussion overdue in the tertiary sector. We provide a detailed explanation of the method, explore the controversies and discuss how systematic review might best be applied within higher education research.

**Literature reviews: themes and variations**

In the higher education literature, the term ‘systematic review’ is used loosely. For example, Konur (2007) refers to a ‘systematic review of four major anti-discriminatory laws’ within a study of disabled students using computer-assisted learning. In this instance, ‘systematic review’ refers simply to a methodical examination of a particular set of documents – such a usage is indicative of the non-technical use of the term.

At the broadest level, a systematic approach to the literature can be distinguished from a narrative review in that it uses a structured system of inquiry to find and review publications. However, the term ‘systematic review’ is increasingly used to refer to a specific type of literature review associated with a specific methodology. Literature reviews can be categorised as either ‘narrative reviews’ or ‘systematic reviews’ (Petticrew, 2001). It may be helpful to think of two types of narrative review. The first is the traditional ‘narrative’ or ‘critical’ literature review, which presents a particular perspective on the literature, framed entirely through the perspective of the author. The second is essentially a narrative approach, with some systematised elements, such as
reproducible searching strategies and a systematic presentation of the studies. These types of narrative literature reviews will be familiar to most readers as these are usually found in theses and less commonly published as papers. Likewise, systematic reviews are not all the same. The standard definition of a systematic review is a literature review that uses a specific methodology to produce a synthesis of available evidence in answer to a focused research question. It is also worth distinguishing the Cochrane (in healthcare) or Campbell (in education, justice and social welfare) systematic review, which are derived from the Cochrane protocol for conducting systematic reviews. The Cochrane/Campbell systematic review uses the general systematic review methodology, but includes a range of other methodological choices, such as independent reviewers and a specific protocol. As these review types are not as well known in higher education, it is worth explaining these methodologies more closely.

The Cochrane handbook for systematic reviews of interventions (Higgins & Green, 2011) outlines key features of all systematic review methodologies. It states that a systematic review has:

- a clearly stated set of objectives with pre-defined eligibility criteria for studies
- an explicit, reproducible methodology
- a systematic search that attempts to identify all studies that would meet the eligibility criteria
- an assessment of the validity of the findings of the included studies, for example through the assessment of risk of bias
- a systematic presentation, and synthesis, of the characteristics and findings of the included studies.

David Gough (2007b), Director of the EPPI-Centre, describes a nine-phase process for systematic reviews:

1. establishing the review question
2. defining inclusion and exclusion criteria
3. articulating the search strategy, including information sources
4. screening the articles to see if they meet the inclusion and exclusion criteria
5. reporting the results of the search strategy, usually through a flowchart
6. extracting relevant data from included studies
7. assessing the methodological quality or rigour of the included studies
8. synthesising, either quantitatively or qualitatively, the collective evidence of the included studies
9. drawing conclusions and communicating these findings in a manner which is relevant to readership.

It is worth noting that the synthesis phase frequently uses formal methods. Meta-analysis is commonly used with quantitative data, particularly experimental data, to statistically estimate the combined result of individual studies. Qualitative synthesis can use a range of different methods such as meta-ethnography, adapted grounded-theory techniques and critical interpretive synthesis (Barnett-Page & Thomas, 2009).

If the above nine phases are part of all systematic reviews, what are the distinguishing features of a Campbell or Cochrane type of systematic review? To answer this question, it is first important to understand the role of these not-for-profit organisations in the development of a culture of systematic review. In healthcare, the Cochrane
Collaboration is dedicated to the conduct of systematic reviews to deliver evidence for healthcare practice and provides a central role in standardising the methodology. In the social sciences, the Campbell Collaboration, an international research network that produces systematic reviews in education, justice and social welfare, has modelled itself on the Cochrane Collaboration. Both organisations are founded on the same principles: working collaboratively, promoting enthusiasm, avoiding duplication, minimising bias, keeping reviews and other outputs up-to-date, relevant and accessible, continuing quality development, provision of continuity of governance, and promoting participation in the collaboration (Campbell Collaboration, n.d.; Cochrane Collaboration, n.d.). The Campbell or Cochrane Collaboration systematic reviews differ from other systematic reviews by:

- expanding the systematic search to unpublished reports to avoid publication bias, such as where results that support the original research hypotheses are more likely to be accepted into peer-reviewed journals (the so-called ‘file drawer problem’)
- having a team that usually spans national boundaries
- following a protocol (project plan) that is developed *a priori* and peer reviewed
- applying inclusion criteria, data extraction and quality assessment that are undertaken independently by at least two reviewers
- being subjected to peer review and editorial review from within the Campbell or Cochrane Collaboration.

See Table 1 for a comparison of systematic and narrative review methodologies. It is worth noting that these are broad categories and individual literature reviews may contain elements of each approach. It may be that for some research questions the systematic review methodology is highly appropriate, while for other questions again, a narrative review may provide better answers. These, and other forms of presenting and synthesising previous literature, can be conceived of as complementary rather than competing.

**Recent use of systematic review in higher education**

The systematic review methodology is difficult to understand in abstract, particularly given the variation in its application. We have provided three examples of relatively recent systematic reviews in higher education within Table 2, describing how each study has addressed each of the nine canonical phases described by Gough (2007b). These examples show both the potential of systematic review (because the synthesised conclusions have value for practitioners and policy-makers alike), as well as the variability of the systematic review methodology.

See Table 3 for the results of a scoping exercise to explore the extent of the systematic review methodology within the higher education sector. Within this brief search of the literature, we used an inclusive definition of ‘systematic review’ to incorporate those reviews that mostly, although not definitively, followed the general nine-phase process, and a definition of ‘higher education’ that excluded the vocational education sector or school student attitudes towards higher education. We searched the ERIC database: on title and abstract, on a range of predominant journals that we believed to be common references for higher education researchers on full text, and on the Campbell Centre and EPPI-Centre web libraries. Once duplicates were removed, we
Table 1. Comparing narrative and systematic literature review types.

<table>
<thead>
<tr>
<th>Review type</th>
<th>Variations (not exclusive)</th>
<th>Phase 1 Phase 2 Phase 3 Phases 4 and 5 Phase 6 Phase 7 Phase 8 and 9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Review question  Inclusion-exclusion criteria</td>
</tr>
<tr>
<td>Narrative</td>
<td>Narrative review</td>
<td>Often general discussion or critical focus</td>
</tr>
<tr>
<td>A literature review</td>
<td></td>
<td>More focused than narrative</td>
</tr>
<tr>
<td>Systematic</td>
<td>‘Non-Cochrane’ systematic review</td>
<td>Clear research question</td>
</tr>
<tr>
<td>Campbell/Cochrane</td>
<td></td>
<td>Clear research question</td>
</tr>
</tbody>
</table>
Table 2. Three systematic reviews in higher education: processes and outcomes across the nine phases of an educational research systematic review methodology.

<table>
<thead>
<tr>
<th>Authors (year) title and source of publication</th>
<th>Phase 1</th>
<th>Phase 2</th>
<th>Phase 3</th>
<th>Phases 4 and 5</th>
<th>Phase 6</th>
<th>Phase 7</th>
<th>Phases 8 and 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>McGrail et al. (2006)</td>
<td>Review question: To review published literature that reports the effectiveness of measures designed to promote publication</td>
<td>Inclusion-exclusion criteria: (1) Reported implementation of a specific structured intervention with the aim of increasing publication rates; (2) The target group were academics or professionals involved in academic work; (3) The article provided data that assessed the effectiveness of the intervention. No exclusions listed.</td>
<td>Search strategy: Databases: Medline, CINAHL, ERIC, PsycINFO and Web of Science, 1984–2004. Search terms provided in general terms. Reference lists of articles also searched.</td>
<td>Screening method and results of search: 17 articles met the criteria. Reasons for exclusion of articles given. No flow chart was provided. No comment regarding use of independent researchers to screen.</td>
<td>Data extraction processes: Full data extraction process not specifically articulated. Two independent reviewers extracted data regarding: number and type of participants; model of intervention (writing support groups, structured writing courses and provision of a writing coach); evaluation methods; impact on publication rate, calculated as publications per person per year, pre- and post-intervention; other impacts.</td>
<td>Assessment of methodological quality and rigour: Quality of papers described narratively, quality assessment scale not specifically provided and methodology not articulated.</td>
<td>Synthesis and conclusions: Described positive effects on publication rates and other positive effects. Concludes that the evidence is limited but that all three models were beneficial, with a trend for support groups to be superior; structured writing courses to be less beneficial.</td>
</tr>
<tr>
<td>Authors (year) title and source of publication</td>
<td>Phase 1</td>
<td>Phase 2</td>
<td>Phase 3</td>
<td>Phases 4 and 5</td>
<td>Phase 6</td>
<td>Phase 7</td>
<td>Phases 8 and 9</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>---------------</td>
<td>--------</td>
<td>--------</td>
<td>--------------</td>
</tr>
<tr>
<td>How effective are ‘virtual patients’ in comparison with no intervention and alternate instructional methods, and what virtual patient design features are associated with higher learning outcomes?</td>
<td>Review question</td>
<td>Inclusion-exclusion criteria</td>
<td>Search strategy</td>
<td>Screening method and results of search</td>
<td>Data extraction processes</td>
<td>Assessment of methodological quality and rigour</td>
<td>Synthesis and conclusions</td>
</tr>
<tr>
<td>Cook, Erwin, &amp; Triola (2010) Computerized virtual patients in health professions education: A systematic review and meta-analysis, Academic Medicine.</td>
<td>Inclusions: (1) Any health profession at any stage of learning. (2) Learners must emulate the roles of health care providers via a computer. Exclusions: Non-interactive cases presented on a computer; simulation delivered by means other than a computer; non clinical simulations.</td>
<td>Worked independently and in duplicate to screen for inclusion. Methodology provided for resolving disagreement. Interrater agreement calculated using intraclass correlation coefficient (ICC). 48 articles from 698 original search yield. Flowchart provided.</td>
<td>Full data extraction only for comparative studies and rigorous qualitative studies. Data abstraction conducted by two or more independent reviewers. ICC calculated. Data included: study type; educational level; topic; aspects of interface and interaction; individual or group learning experience; type of outcome evaluated. A range of further data were extracted as relevant, e.g. comparison type, themes for qualitative studies.</td>
<td>Quality assessment procedure described for comparative studies but not for ‘rigorous’ qualitative studies. Quality indicators included: representativeness of the intervention group; selection of the comparison group; and comparability of cohorts.</td>
<td>Meta analysis of comparative studies and pooling of themes from qualitative studies. Pooled effect sizes (ES) for studies of virtual patients compared with no interventions were large (&gt;= 0.8) for a range earning outcomes and confidence intervals (CI) excluding small effects (&lt; 0.5). Pooled ES for studies of virtual patients in comparison with non-computer instruction were small (0.17–0.10) with non-significant CI. Common themes identified from qualitative studies point to common experiences and possible design considerations.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Authors (year) title and source of publication</td>
<td>Phase 1</td>
<td>Phase 2</td>
<td>Phase 3</td>
<td>Phases 4 and 5</td>
<td>Phase 6</td>
<td>Phase 7</td>
<td>Phases 8 and 9</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------------</td>
<td>---------</td>
<td>---------</td>
<td>---------------</td>
</tr>
<tr>
<td><strong>Canadian Council on Learning (2006)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>How is quality in post-secondary education measured?</em></td>
<td><em>How is quality in post-secondary education measured?</em></td>
<td><em>How is quality in post-secondary education measured?</em></td>
<td><em>Inclusions: (1) Post 1985, post-secondary education in OECD.</em></td>
<td><em>(2) Studies must identify the indicators used to measure quality or quality indicators and analyse the tools or methods used to measure quality of post-secondary education as a whole.</em></td>
<td><em>Exclusions not listed.</em></td>
<td><em>Databases: Academic Search Premier,</em></td>
<td><em>Developed a scoring rubric,</em></td>
</tr>
<tr>
<td><em>Internet.</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>Synthesis takes the form of presenting the commonly considered indicators.</em></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Authors (year) title and source of publication</strong></th>
<th><strong>Phase 1</strong></th>
<th><strong>Phase 2</strong></th>
<th><strong>Phase 3</strong></th>
<th><strong>Phases 4 and 5</strong></th>
<th><strong>Phase 6</strong></th>
<th><strong>Phases 7</strong></th>
<th><strong>Phases 8 and 9</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Canadian Council on Learning (2006)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>How is quality in post-secondary education measured?</em></td>
<td><em>How is quality in post-secondary education measured?</em></td>
<td><em>How is quality in post-secondary education measured?</em></td>
<td><em>Inclusions: (1) Post 1985, post-secondary education in OECD.</em></td>
<td><em>(2) Studies must identify the indicators used to measure quality or quality indicators and analyse the tools or methods used to measure quality of post-secondary education as a whole.</em></td>
<td><em>Exclusions not listed.</em></td>
<td><em>Databases: Academic Search Premier,</em></td>
<td><em>Developed a scoring rubric,</em></td>
</tr>
<tr>
<td><em>Internet.</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><em>Synthesis takes the form of presenting the commonly considered indicators.</em></td>
</tr>
</tbody>
</table>
Table 3. Scoping search strategies for systematic reviews in higher education in peer-reviewed journals or in recognised systematic review centres (November 2011).

<table>
<thead>
<tr>
<th>Source</th>
<th>Systematic review search terms</th>
<th>Higher education-related search terms</th>
<th>Total search yields</th>
<th>Number of systematic reviews in higher education excluding health professions</th>
<th>Number of systematic reviews from health professions or vocational education</th>
<th>Number of publications that were not systematic reviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERIC database</td>
<td>‘Systematic review’</td>
<td>‘Higher education’</td>
<td>16</td>
<td>3</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>Higher Education Research &amp; Development</td>
<td>As above</td>
<td>n/a</td>
<td>5</td>
<td>1</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Journal of Higher Education</td>
<td>As above</td>
<td>n/a</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Review of Higher Education</td>
<td>As above</td>
<td>n/a</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Research in Higher Education</td>
<td>As above</td>
<td>n/a</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Quality in Higher Education</td>
<td>As above</td>
<td>n/a</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>International Journal of Educational Research</td>
<td>As above</td>
<td>n/a</td>
<td>7</td>
<td>0</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>Studies in Higher Education</td>
<td>As above</td>
<td>n/a</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Assessment and Evaluation in Higher Education</td>
<td>As above</td>
<td>n/a</td>
<td>4</td>
<td>0</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Educational Researcher</td>
<td>As above</td>
<td>n/a</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Campbell Collaboration</td>
<td>n/a</td>
<td>Reviews from Education Group</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>EPPI Centre</td>
<td>n/a</td>
<td>Any education-related review</td>
<td>16</td>
<td>2</td>
<td>5</td>
<td>9</td>
</tr>
</tbody>
</table>
found five systematic reviews in higher education. These were in disparate areas of study and included reviews of: interventions to increase academic publication rates (McGrail et al., 2006), interdisciplinary higher education (Spelt et al., 2009), links between administration and student outcomes (Jackson & Kile, 2004), professional identity development (Trede et al., 2012), work-based learning (Scesa & Williams, 2008) and personal development planning (Gough et al., 2003). This is far less than the number of educational systematic reviews of school education on the EPPI Centre website or the number of health professional education systematic reviews on the BEME website. Only one of our example systematic reviews in Table 2 is drawn from this list, the other two are taken from the health professional education literature and the ‘grey’ (not readily sourced via academic publishers) literature, respectively.

The value of systematic review methodology to educational research

Given the lack of penetration of systematic reviews into the higher education literature, it is important to ask what benefits there are in adhering to such a rigorous and time-consuming literature review methodology. As the examples in Table 2 indicate, systematic review allows the reader to make their own judgements as to the quality and meaning of the evidence, as well as a judgement on the quality of the review methodology itself. The breadth of the search allows for all research to be considered, not just the most accessed or best known work. In this way, an educator or educational researcher can quickly scan what is known about a topic, how trustworthy the research conclusions are and what the collective implications of the research are. A good systematic review saves reproduction of literature searching, directs readers to quality literature and provides a formal synthesis of the research outputs. Gough (2007a) summarises the key benefits of systematic review to educational policy and practice as:

- **access** to the literature, with its ever increasing publication base
- **providing trustworthiness and accountability** of the literature review (Gough argues that the methods of systematic review allow us to ‘be clear about what we know and how we know it within different ideological and theoretical positions’ (p. 35)
- **revealing conceptual and value positions**, through mechanisms designed to make explicit the inevitable biases, which might otherwise be ‘hidden within the discourse of the account of knowledge’ (p. 35).

Critiques of systematic review within the educational research community

In contrast to the healthcare education sector, systematic review has been heavily critiqued by some sectors of the educational research community. It is important to examine these debates, as they likely influence the use of the systematic review methodology within the tertiary sector. We argue that the majority of concerns revolve around three interrelated themes:

1. Positioning of the positivist paradigm. There is a body of work arguing that systematic reviews privilege certain kinds of research paradigms, either explicitly or through the language employed to describe them (e.g. Clegg, 2005; Denzin, 2009; Hammersley, 2001; Suri & Clarke, 2009). Associated with this is the
elevation of the systematic review over other sorts of literature reviews (Hammersley, 2001; Suri & Clarke, 2009). For example, Suri and Clarke (2009) write: ‘The rhetorical effect of terms such as evidence-based practice, systematic reviews, clarity, comprehensive, reliable, objectivity, and replicable not only discredits any opposition but also has the political impact of favouring positivism’ (p. 400). These arguments stem from the notion that systematic reviews, like any other form of research practice, are ideological in conception and in practice. Denzin (2009) explicitly links systematic review to the rise of neoliberal regulatory government processes and their associated emphasis through performance measurement.

(2) Adaptation of methodological choices to a social sciences context. Within the educational research literature concerned with the elevation of systematic review and its underpinning ‘measurement’ principles, there is an associated unease regarding the lack of critical interrogation of systematic review methodology. In this argument, the emphasis on establishing a priori criteria for the inclusion of studies, the transparency of method and the emphasis on objectivity may lead to what Maclure (2005) terms ‘clarity bordering on stupidity’ (p. 1). In particular, there is a concern that forms of research that are heavily qualitative and descriptive in nature do not mesh with underlying principles of systematic review. Qualitative synthesis methodologies within a systematic review must take account of highly divergent traditions, such as phenomenography, ethnography, discourse analysis, phenomenology and so on. There is a range of publications that note that formal qualitative synthesis will be challenged by variant analysis methods and presentation forms in order to transform what is essentially context-specific information into a more generalisable form (Barnett-Page & Thomas, 2009; Ring, Ritchie, Mandala, & Jepson, 2010).

(3) Practical applicability to policy and practice. Clegg (2005) describes, in the higher education sector, the limited use of the conclusions of systematic review. She illustrates this with the example of ‘the sense of lack’ for practitioner use with respect to a systematic review of personal development planning in higher education (Gough et al., 2003). Clegg (2005) argues that a focus on outcomes means that the review does not provide sufficient information for educationalists regarding particular mechanisms of what works and when.

These debates are important to acknowledge, not only to avoid the rhetorical positioning of ‘systematic review’ as unquestioningly superior, but also to avoid throwing the proverbial baby out with the bathwater. There is no question that systematic review derives from a strongly positivist paradigm, with its foundations in meta-analysis of randomised controlled trials (Cochrane Collaboration, n.d.), but it is another step altogether, therefore, to discredit the process. Of course, the unquestioning acceptance of any methodology is not good academic practice and it is important to critically interrogate methodologies of systematic review and to understand the arguments surrounding synthesis of divergent studies (Andrews & Harlen, 2006). This occurs also within the healthcare systematic review community, where using quality appraisal scales with qualitative research is debated (Dixon-Woods et al., 2006) or, in some situations, a focus on processes rather than outcomes is suggested (Pawson, Greenhalgh, Harvey, & Walshe, 2005). In all instances, the concept of evidence does not exist in isolation within an ivory tower, but has broad societal and political implications.
Potential use of systematic reviews in higher education

The broader question of the applicability of systematic review to higher education is a complex one because of the multi-disciplinary nature of this broad field. In exploring the potential use of systematic review in higher education research, several issues will need to be considered, including sectoral readiness, diversity of research traditions, the feasibility of the uptake of the processes involved and the likely benefits of uptake. We will use the example of our own work to provide context to these matters.

Illustrative example: measuring teaching quality

Our research team has commenced a Campbell/Cochrane-style systematic review in the area of measuring teaching quality in higher education. This review is designed to answer the question: what is known about the measurement of teaching quality? Our primary aims are: to define teaching quality and identify the common constructs associated with teaching quality, to establish how teaching quality is typically measured or assessed, to establish and evaluate what is known about reliability of available measures/instruments, and to recommend which one instrument or set of measures to use or whether to develop our own. In line with the systematic review process, we have pre-defined criteria for inclusion. Examples of our criteria are: all included studies must report on instruments, or the development of instruments, that measure, assess, appraise, evaluate or analyse teaching, and all included studies must report in any tertiary teaching domain. We have established comprehensive search terms, including common synonyms for the terms measurement, instrument, teaching and higher education, and used these to search 13 databases, including ERIC, PsychINFO, British Education Index, Current Contents, Expanded Academic, Proquest Education Journals, A+ Education, CBCA Education, Campbell Collaboration, Medline, CINAHL and PubMed, together with handsearching of reference lists and contact with content experts. The resulting 11,173 articles were reduced through scan of title to remove obviously unrelated publications, with the remainder independently reviewed by two researchers to see if they meet the set criteria, with disagreements resolved by consensus or additional reviewers. This time-consuming and labour-intensive phase of the review has resulted in 47 included articles. We are in the process of extracting the data from these articles and establishing their methodological quality using standardised descriptors, with each process again performed by two independent researchers. Undertaking such an intensive collaborative exercise has enlightened us to many of the issues surrounding systematic review, which we discuss in the following sections.

Sectoral compatibility and readiness

Systematic review may be easily acceptable within medical education because the approaches to research that are most aligned with systematic review epistemologies are found within medical research. Medical research tends towards a more uniform approach to research, in both method and approach to reporting. But educational research celebrates a range of approaches, more commensurate with its diverse traditions in critical, feminist, post-structuralist, post-modern, hermeneutic, interpretivist and other post-positivist approaches to knowledge. This is particularly true for higher education, where the approaches to educational research may be strongly influenced by various disciplinary research cultures.
Furthermore, for systematic review to be useful in the higher education sector, it has to be applicable to the literature in the sector. For example, theory-based publication simply does not suit systematic review. This is also a matter of philosophical alignment, in that practitioners and researchers within the sector must wish to base decisions upon formal synthesis of a broad range of evidence. In our view, there are subsectors within higher education which have both research that is suited to systematic review and practitioners who are open to using the results of systematic reviews.

We return to our example of measuring teaching quality. Preliminary results indicate a range of appropriate studies, both quantitative and qualitative, that provide empirical investigations of instruments that measure teaching quality. The topic itself is contentious in nature, with political as well as practice implications. This, in our view, is a good area for systematic review, where the collected evidence can provide valuable information to on-the-ground educators as well as policy makers.

Managing diversity of research traditions

As stated before, the higher education sector is multi-disciplinary and lacks a single unifying paradigm that underpins research and reporting. At various junctures in the systematic review process this can lead to challenging issues. In developing the question and search terms, for instance, there may be different understandings of the meanings of the words. For example, in our experience, not all members of our systematic review team interpreted the word ‘instrument’ in the same manner. In exactly the same way, when searching the literature databases, terminology may be highly diverse and specific keywords may not be mentioned in title, abstract or, indeed, the entire text. The prior established criteria for inclusion and exclusion of papers are probably more likely to result in false positives and false negatives when there is ambiguity surrounding key terms. As previously mentioned, synthesis must work with a wide variance in study methodologies, reporting standards and epistemologies. Due to this variance, a team approach to systematic review is critical within higher education, along with the use of moderation to arrive at shared understandings of inclusion/exclusion criteria.

Feasibility of processes

The systematic review processes themselves do not pose any impediment to their application in higher education, although they are time intensive. Our experience indicates that the moderation of the selection rules is easily achieved with a small number of colleagues. If the criteria and their application are unambiguous to the researchers and the readership, the whole process will flow relatively efficiently. This highlights the importance of the process in establishing the selection rules and other analysis criteria. In our experience, however, with many of the team unfamiliar with the systematic review methodology, training becomes critical.

Although the processes themselves can be followed easily enough, the required time investments could be a barrier to the feasibility of systematic review being routinely used in higher education, especially where the process is not adequately supported with funding or where expert guidance is not readily available.

Benefits

We believe that a growing awareness and use of systematic review will be an asset to the higher education sector. The three examples provided earlier in this paper (Table 2)
provide meaningful conclusions to policymakers and practitioners. McGrail et al. (2006) provide clear guidance for academics wishing to enhance publication, Cook et al. (2010) clearly indicate that virtual patients are of educational benefit, and indicate likely successful design choices for this type of computer-assisted learning, while the Canadian Council for Learning’s (2006) report provides recommendations for universities wishing to measure the quality of their institution in meaningful ways. As alluded to earlier, there are specific subsectors within higher education that would particularly benefit from systematic review.

In health professional education, another benefit of systematic reviews has been the forming of international collaborations and partnership through large review teams in the BEME organisation. As with the Cochrane and Campbell Collaborations, these research networks are inclusive and assist in dissemination of research results through engagement of reviewers. More than that, the work helps define the discipline. For example, BEME’s reviews are on self-assessment, student portfolios, feedback, faculty development, inter-professional learning and educational games (BEME, n.d.). These are all key topics within health professional education and reviewers are drawn from all over the globe. It is appealing to consider the value of such an organisation exclusively focused upon higher education.

One consequence of the emerging interest in the conduct of systematic review in higher education might be the enhancement of the quality of research and reporting. Systematic reviews can build upon the current work of educational researchers through identifying specific strengths and weaknesses in the literature. Systematic review may also facilitate researchers and practitioners in simultaneously understanding and influencing the overall practice and status of higher education research.

**Conclusion**

Systematic review is a specific, carefully defined approach to the literature review, which has not reached its full potential in its application to higher education. Indeed, this article is addressing the possibility that many within the higher education sector are unaware of the precise detail of the systematic review process. The value of systematic reviews is in providing a transparent, comprehensive and structured approach to searching, selecting and synthesising the literature. There are some caveats surrounding application of the methodology to the diverse traditions within higher education, which is a particularly complex overlap of a variety of disciplines. Systematic reviews should be applied only when they can provide a valid means to summarise the literature and it should be acknowledged that this may not be the case for all areas of higher education research. We propose that the systematic review methodology, associated with a culture of collaboration, rigour and methodological debate, has potential for the higher education sector. We believe that systematic review can assist us in making choices that will assist our students to learn because of, not despite, our efforts.

**Acknowledgements**

The authors wish to acknowledge the insightful comments made by Professor Richard James and Dr Phillip Dawson on earlier versions of this paper. This work has been supported by the Australian Learning and Teaching Council through the strategic priority projects program (SP10-1845).
References


Dixon-Woods, M., Bonas, S., Booth, A., Jones, D., Miller, T., Shaw, R., … Young, B. (2006). How can systematic reviews incorporate qualitative research A critical perspective. *Qualitative Research, 6*(1), 27–44.


