



Contents lists available at ScienceDirect

Environment International

journal homepage: [www.elsevier.com/locate/envint](http://www.elsevier.com/locate/envint)

## “A little learning is a dangerous thing”: A call for better understanding of the term ‘systematic review’

Neal R Haddaway\*, Magnus Land, Biljana Macura

Mistra EviEM, Stockholm Environment Institute, Box 24218, 104 51 Stockholm, Sweden

### ARTICLE INFO

#### Article history:

Received 6 September 2016

Received in revised form 20 December 2016

Accepted 21 December 2016

Available online xxx

#### Keywords:

Evidence-based policy

Evidence

Evidence synthesis

Grey literature

Risk of bias

Science policy

### ABSTRACT

Systematic reviews are becoming a widely accepted *gold standard* in evidence synthesis for evidence-based and –informed policy and practice. Many organisations exist to coordinate the registration, conduct and publication of systematic reviews across a range of disciplines, including medicine, international development, and environmental management and biodiversity conservation. As the term ‘systematic review’ becomes more widely recognised, however, there is a risk that stakeholders may have only partial understanding of the rigorous methods required to produce a reliable systematic review. Here, we highlight one such example from the field of education and international development, where a World Bank report claimed to ‘systematically review’ six ‘systematic reviews’ that found divergent results. We critically appraise the six included reviews and the World Bank report itself using an a priori quality assessment tool. Our analysis shows that none of the six included reviews are classifiable as systematic reviews according to widely accepted criteria. We also find that the World Bank report failed to use true systematic review methods to synthesise the included reviews findings. Our study demonstrates the risks associated with partial understanding of the added value associated with systematic reviews and highlights a need for improved awareness of what systematic reviews are.

© 2016 Elsevier Ltd. All rights reserved.

### 1. Introduction

Systematic reviews aim to provide a *gold standard* in summarising documented scientific evidence (Pullin & Stewart, 2006). They typically involve the following step-wise processes: setting out the planned methods in a peer-reviewed and published protocol; searching for evidence (including grey literature in an attempt to minimise publication bias); assessment of the relevance and validity of each piece of evidence; extraction of study descriptors and findings; and, synthesis and reporting of the evidence base identified. Throughout the process reviewers attempt to maximise comprehensiveness, transparency, repeatability and objectivity.

Several organisations that coordinate systematic reviews (The Cochrane Collaboration in medicine and health science (The Cochrane Collaboration, 2016); The Campbell Collaboration (The Campbell Collaboration, 2016) and the EPPI Centre (The EPPI-Centre, 2016) in social science; and The Collaboration for Environmental Evidence in environmental management (The Collaboration for Environmental Evidence, 2016)) act as review coordinating bodies, publishing protocols and final review reports following a thorough peer-review, thereby endorsing the reviews and ensuring a high standard of conduct according to established guidelines (e.g. Higgins & Green, 2011). These aim to

establish best practice and minimum standards for systematic reviews across a range of different disciplines, as introduced above. To maximise reliability of reviews, the organisations have strict minimum requirements that must be assessed through external peer-review and met before they will publish a systematic review (or review protocol) (e.g. Higgins & Green, 2011; Steering Group of the Campbell Collaboration, 2015; The Collaboration for Environmental Evidence, 2013).

As a result of their rigorous methods, systematic reviews are regarded as authoritative and repeatable (Haddaway & Pullin, 2014) and are viewed as the most reliable of evidence synthesis methods across a range of disciplines (Petticrew, 2001).

Systematic review methods are now ubiquitous within some disciplines such as medicine. However, the methods are still relatively novel in other fields, such as environmental management and education (Pullin & Stewart, 2006). Thus, whilst the number of stakeholders aware of the term ‘systematic review’ increases promisingly, there is a lag between those aware of the term and those who appreciate the steps necessary to make a systematic review reliable.

Here, we outline a recent example from the field of education and international development where a high-level, international organisation regrettably misunderstood the term ‘systematic review’ (Evans & Popova, 2015a), failing to include the necessary rigour in their own review that warrants such a label, and misidentifying non-systematic reviews as systematic reviews. This case highlights the dangers of improper awareness of systematic review methods and prompts a call

\* Corresponding author.

E-mail address: [neal\\_haddaway@hotmail.com](mailto:neal_haddaway@hotmail.com) (N.R. Haddaway).

for raised awareness in many stakeholder groups, including research funders, practitioners, and policy-makers.

## 2. The problem

Review commissioners and end-users of reviews sometimes wonder whether parallel systematic review teams would arrive at the same conclusion, a claim often made by the review coordinating bodies above. Methodologists have previously considered this problem (e.g. Hopyan & Mugford, 1999; Jadad et al., 1997), and have carefully examined differences in methodological approaches taken by reviewers to ascertain why conclusions may differ (Thompson et al., 2008). A recent World Bank report (Evans & Popova, 2015a) and accompanying blog (Evans & Popova, 2015b) claimed to have found evidence of 'identical' systematic reviews that had been completed independently, which had arrived at divergent conclusions. Such a finding, the authors say, highlights a fundamental flaw with systematic review methodology.

In the blog (Evans & Popova, 2015b) based on a World Bank research report (Evans & Popova, 2015a) assessing what works in improving learning outcomes in developing countries, the authors claim to identify six systematic reviews and, in comparing the reviews' findings, fail to discover much overlap between the review conclusions. In further analysis, they discover that this divergence in findings is mainly driven by differences in the included primary evidence and the formulation of intervention categories. Due to the contradictions in the findings of what are perceived as similar systematic reviews, the authors rightfully ask 'how definitive are these systematic reviews really?' and caution that the community should 'take systematic reviews with a grain of salt'.

In this response, we argue that the authors of the mentioned report are at risk of constructing a 'straw man argument', since neither their own review nor the majority of the reviewed reviews can be considered as true systematic reviews. Our argument is based on a rigorous assessment of both the authors' review of reviews, and the systematic reviews it included. Our assessment used a structured critical appraisal tool applied by three independent reviewers, which is explained in more detail below. Finally, we outline a number of lessons from this example and stress the need for improved awareness of systematic review methods.

## 3. Methods

Whilst no universal definition of a 'systematic review' exists across disciplines, three broad minimum standards are common to all systematic reviews: i) systematic reviews' methods should be described in sufficient detail to allow full repeatability and traceability; ii) they must include a systematic approach to identifying and screening relevant academic and grey literature, iii) they should include critical appraisal of the validity (internal, i.e. quality, and external, i.e. generalisability) of included studies to give greater weight to more reliable studies (Higgins & Green, 2011; Steering Group of the Campbell Collaboration, 2015; The Collaboration for Environmental Evidence, 2013). We have used these minimum standards to produce a schema that aided our critical appraisal of the reviews that were included in Evans and Popova's review (Evans & Popova, 2015a). Other authors have previously produced appraisal criteria for scoring different types of evidence review based on the quality of their methods (Woodcock et al., 2014). However, we have chosen to use a specifically designed tool that qualitatively appraises the methods used in each review, rather

**Table 1**  
Schema used to critically appraise studies included in Evans and Popova (Evans & Popova, 2015a).

Domain	Questions	Explanation
Nomenclature	Does the document refer to itself as a "systematic review"? How does the document refer to itself? Is the review published in an academic journal? Was the review subject to peer-review? If so, how (external/internal)?	Reviews may not claim to be systematic. Peer-review (whether formal or informal) is a central cornerstone of scientific research and indicates some form of community appraisal.
Protocol	Was a protocol produced (as mentioned in the review)? Was this protocol externally peer-reviewed? Was the protocol published?	A review protocol sets out the methods for the review and allows expert and public input to fine-tune the sources and strategies for identifying and including the best available evidence.
Searching	Were multiple academic sources searched? Was a search string established and used in all resources? Were searches documented (minimum date, search terms, numbers of results)? Were attempts to search for grey literature included?	Systematic reviews should be as comprehensive as possible, searching multiple databases and making efforts to search for grey literature in addition to published research. Search activities should be documented in sufficient detail to allow the review to be repeated.
Screening	Are clear inclusion criteria reported? Was screening undertaken by multiple reviewers to any extent? Was consistency between reviewers tested? Are the results of screening (numbers) reported for titles and abstracts? Are the results of screening (numbers and exclusion reasons) reported for full texts? Are the number of unobtainable/untranslated articles reported?	Transparent criteria for the inclusion of articles in the review are vital to demonstrate objectivity and allow repeatability in a review. Screening should be carried out by more than one reviewer to demonstrate objectivity and consistency should be assessed between reviewers. Screening activities should be clearly documented for traceability. Numbers of articles making it through each stage should be documented. Ideally, reasons for excluding articles at full text should be provided.
Critical appraisal	Are included studies appraised for internal and external validity? Are the criteria for CA provided in detail? Are the results of CA reported in detail?	Critical appraisal of study internal (quality) and external (generalisability) validity is a vital stage in every systematic review. Synthesis should be based on reliable evidence. Critical appraisal allows for unreliable evidence to be excluded or down-weighted in analyses. Activities should be documented and decisions should be justified in detail.
Data extraction	Is the method for data extraction reported in detail? Are the extracted data reported? Is any data manipulation reported in detail?	Data (quantitative and qualitative) should be extracted in a transparent way to ensure that objectivity and consistency across studies is maintained. Any manipulation of data (e.g. calculation of effect sizes) should be documented in detail. Ideally, all extracted data should be available in some form.
Synthesis	Is a narrative/qualitative/quantitative synthesis present? Is there any evidence of vote-counting? Is publication bias assessed or discussed? Are synthesis methods provided in detail? Are synthesis outputs reported in detail?	Narrative synthesis is the discussion of the evidence base as a whole. Qualitative synthesis involves an established means of combining results in a qualitative way. Quantitative synthesis involves the use of powerful statistics. Vote-counting should always be avoided since it ignores effect sizes and can mask underlying patterns in the data that are not apparent in individual studies. Publication bias should always be investigated to assess whether the findings of the review may be affected by the lack of studies with certain findings (e.g. non-significant or contradictory results). All synthesis activities should be documented in detail.
Other	Are potential conflicts of interest discussed? Is it obvious who funded the review?	Potential conflicts of interest should be dealt with in an acknowledgements section and through documenting author affiliations. Sources of funding should be included in the acknowledgements.

Download English Version:

<https://daneshyari.com/en/article/5748428>

Download Persian Version:

<https://daneshyari.com/article/5748428>

[Daneshyari.com](https://daneshyari.com)