Environment International xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Environment International

journal homepage: 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 xxxx

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.

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 (N.R. Haddaway).

http://dx.doi.org/10.1016/j.envint.2016.12.020
0160-4120/

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. Hopayian & 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.

Please cite this article as: Haddaway, N.R., et al., "A little learning is a dangerous thing": A call for better understanding of the term 'systematic review', Environ Int (2016), http://dx.doi.org/10.1016/j.envint.2016.12.020

than relying on quantitative scoring tools that rank reviews but may miss valuable descriptive information explaining the different approaches.

We use 8 domains to assess how well the six reviews included in the Evans and Popova report conform to systematic review standards: nomenclature, protocol, searching, screening, critical appraisal, data extraction, synthesis and other. A detailed description of each domain and the questions that were used as prompts to complete the assessment are provided in Table 1. Each of the six review articles considered in the report were critically appraised by one reviewer with two additional reviewers double-checking each assessment. Any disagreements were discussed as a group and the appraisal modified to reflect group consensus. We used the same tool to critically appraise the report by Evans and Popova itself that was also reviewed independently by all three reviewers, and all three assessments are provided in the results for clarity.

4. Results

4.1. Critical appraisal of the studies included in the World Bank report

The results of our critical appraisal of the six studies in the report and the review report itself (Evans & Popova, 2015a) are summarised in Tables 2 and 3, and provided in full in Additional File 1.

None of the six included studies claim to be systematic reviews. As Table 2 shows, two of the reviews purport to be 'meta-analyses' (Conn, 2014; McEwan, 2014), whilst the rest claim to be 'reviews' (Glewwe et al., 2014; Kremer et al., 2013; Krishnaratne et al., 2013; Murnane & Ganimian, 2014). One review (Krishnaratne et al., 2013) is an update of a published 3ie systematic review, but the update is not carried out in a systematic way, with a lack of a systematic search and no methodological details, which together significantly reduce its reliability. Two of the reviews were not peer-reviewed (Krishnaratne et al., 2013; Murnane & Ganimian, 2014): not itself an explicit indication of poor quality, but an indication of high potential risk of unreliability or bias.

Glewwe et al. (2014) uses systematic approaches for a proportion of the review's design, but the lack of methodological detail and reporting of results mean that it is of highly questionable reliability (Table 2). McEwan's (2014) review is a sophisticated meta-analysis that includes some systematic approaches to searching, and could perhaps be referred to as a systematic review, although there is no critical appraisal, which is a compulsory step for any review published by one of the review coordinating bodies described herein. The review conducted by Conn (2014) appears to be a high quality review and claims to include a systematic literature search but fails only on its lack of methodological detail and documenting of activities.

4.2. Critical appraisal of the World Bank report

Since thousands of systematic reviews are now published, and overlap or replication within a specific subject narrow topic is common, 'tertiary' reviews (systematic reviews of systematic reviews) are becoming common (especially in clinical medicine) (e.g. Torgerson, 2007). In such reviews the rigorous methodological standards are maintained and key principles in methodology are the same as for 'secondary' (systematic) reviews. The blog post (Evans & Popova, 2015b) that accompanies the World Bank report (Evans & Popova, 2015a) cannot be classed as a tertiary systematic review, however, since tertiary systematic reviews must follow the same rigorous systematic review methodology as secondary systematic reviews.

Table 2
Summary of critical appraisal of six reviews in Evans & Popova (2015a).

		Conn (2014)	Glewwe et al. (2014)	Kremer et al. (2013)	Krishnaratne et al. (2013)	McEwan (2014)	Murnane & Ganimian (2014)
Nomenclature	Does the document refer to itself as a "systematic review"?	No	No	No	Yes	No	No
	Was the review subject to peer-review? If so, how (external/internal)?	Yes	Unclear	Yes	No	Yes	No
Protocol	Was a protocol produced? (as mentioned in the review)	No	No	No	No	No	No
	Was this protocol externally peer-reviewed?	n/a	n/a	n/a	n/a	n/a	n/a
	Was the protocol published?	n/a	n/a	n/a	n/a	n/a	n/a
Searching	Were multiple academic sources searched?	Yes	Yes	Unclear	Yes	No	Unclear
	Was a search string established and used in all resources?	Yes	Unclear	Unclear	Yes	No	No
	Were searches documented (minimum date, search terms, numbers of results)	Yes	Unclear	No	No	Partial	No
	Were attempts to search for grey literature included?	Yes	Partial	Unclear	Yes	Yes	No
Screening	Are clear inclusion criteria reported?	Yes	Unclear	Partial	Yes	Partial	Partial
	Was screening undertaken by multiple reviewers to any extent?	No	Yes	Unclear	Unclear	No	No
	Was consistency between reviewers tested?	No	No	Unclear	Unclear	No	No
	Are the results of screening (numbers) reported for titles and abstracts?	Partial	Partial	No	No	No	No
	Are the results of screening (numbers and exclusion reasons) reported for full texts?	No	No	No	No	Partial	No
	Are the number of unobtainable/untranslated articles reported?	No	No	No	No	No	No
Critical appraisal	Are included studies appraised for internal and external validity?	Yes	Partial	No	Unclear	No	No
	Are the criteria for CA provided in detail?	Yes	Yes	No	No	No	No
	Are the results of CA reported in detail?	Yes	No	No	No	No	No
Data extraction	Is the method for data extraction reported in detail?	Yes	No	No	No	Yes	No
	Are the extracted data reported?	No	No	No	No	Yes	No
	Is any data manipulation reported in detail?	No	n/a	No	No	No	No
Synthesis	Is a narrative synthesis present?	Yes	Yes	No	Yes	No	Yes
	Is a qualitative synthesis present?	No	No	Partial	Yes	No	No
	Is a quantitative synthesis present?	Yes	No	Yes	Yes	Yes	No
	Is there any evidence of vote-counting?	No	Yes	Partial	No	No	Yes
	Is publication bias assessed or discussed?	Yes	No	No	No	No	No
	Are synthesis methods provided in detail?	Yes	No	No	No	Yes	No
	Are synthesis outputs reported in detail?	Yes	No	No	Yes	Yes	No
Other	Are potential conflicts of interest discussed?	No	No	No	No	No	No
	Is it obvious who funded the review?	No	No	No	No	Yes	Yes

Green indicates a fulfilled characteristic, red a missing characteristic, orange a partially fulfilled characteristic and turquoise unclear.

Table 3
Summary statements describing the six reviews in Evans & Popova (2015a).

	High quality review, but undertaken by only one reviewer, so opportunity for subjective decisions is high. Incomplete documentation of activities reduces repeatability and traceability.
Conn (2014)	Only partial critical appraisal undertaken. Low level of detail in documentation of activities. Vote-counting ignores magnitude of study effects and ignores patterns across studies.
Glewwe et al. (2014)	No critical appraisal undertaken. Very low level of methodological detail and results not well-documented. Repeatability and traceability accordingly very low.
Kremer et al. (2013)	Review based on a partial update of a 3ie systematic review. The original systematic review was peer-reviewed and endorsed, but the update appears to include a non-systematic and incomplete search with a low level of methodological detail.
Krishnaratne et al. (2013)	A sophisticated meta-analysis that includes a partial systematic search, favouring snowballing over searches of academic databases. No critical appraisal is involved, and strict exclusion criteria relating to effect sizes and statistics may have introduced bias.
McEwan (2014)	Extremely limited methodology. No critical appraisal. Activities not reported in any detail. Not peer-reviewed.
Murnane & Ganimian (2014)	

The methods used are particularly deficient, precluding repeatability. The evidence obtained for the review is not the result of a systematic search: only one database of systematic reviews was searched and no details of the search strategy are given. Fairly detailed inclusion criteria are given, although the means by which these criteria were applied are not reported at all. No critical appraisal has been conducted, explaining the inclusion of several reviews that are not systematic reviews. Neither are methods provided for how data were extracted and synthesised, and it is unclear whether the data reported are all that were available from each study. The review itself is, in essence, a vote-counting exercise ignoring effect size and reliability (The Collaboration for Environmental Evidence, 2013) and no formalised qualitative or quantitative synthesis are undertaken.

5. Discussion

None of the reviews included in Evans & Popova's, (2015a) report can be classified as true systematic reviews. Moreover, the way in which the review of reviews was undertaken is not as rigorous, transparent and replicable as one would expect a systematic review of reviews should be. Nevertheless we can draw a number of lessons from this exchange that can be beneficial for both the synthesis community and end-users.

5.1. The need for improved awareness of systematic review methods

There seems to be an underlying belief throughout Evan and Popova's report and blog that all reviews which follow some form of methodological system should automatically be counted as systematic reviews. This perception is mistaken. A review may include a systematic search or screening, but unless it includes all of the aspects of a full systematic review, such as critical appraisal and full transparency, the review reliability is reduced and it cannot be referred to as systematic. Counting such reviews as systematic reviews in Evans and Popova's report illustrates a confusion that is made by many other authors. It is unhelpful to classify "narrative reviews" as systematic reviews, as the authors do in their paper: whilst narrative synthesis is a formal and acknowledged component of a systematic review (and may be all that is feasible where neither qualitative nor quantitative synthesis is possible), there is no room within the systematic review taxonomy for products that use exclusively narrative, opinion-based approaches to synthesis. Such a review, that follows a narrative thread informed by expert opinion rather than following a set of a priori, established systematic methods, necessarily remains a literature review. **Having a system to search for studies or to define inclusion criteria does not mean that the review can be classified as a systematic review**.

5.2. The benefits of endorsement

Systematic review coordinating bodies are an invaluable asset to the field of research synthesis. They help to prevent – to cite one of Evans and Popova's concerns – "reviews [from varying] in how systematically they define the strategy used to identify the papers reviewed" by using a uniform system to report search results and inclusion criteria (that is, the PRISMA diagram and the commonly used PICO criteria (Moher et al., 2009; The Collaboration for Environmental Evidence, 2013)). Had the reviews within Evans and Popova's study been registered with a systematic review coordinating body such as the Collaboration for Environmental Evidence (The Collaboration for Environmental Evidence, 2016), the authors would have had access to a range of methodological experts and resources, including detailed methodological guidelines and expert advice through peer-review. There would have been an associated requirement to make their review methodology public a priori, subjecting their search strategy, inclusion criteria, and synthesis methods to peer-review. This would have exposed fundamental flaws in the review products.

In the long-run, the term 'systematic review' in a variety of disciplines alongside medicine should be associated with certain minimum standards such as those set out by these coordinating bodies. In this way, the reliability of a review product will be clearly guaranteed through 'endorsement'.

6. Conclusions

Since at least two of the reviews included by Evans and Popova are simply traditional, narrative literature reviews, we should therefore not be surprised, that each of these reviews yield different conclusions. Furthermore, Evans and Popova do not consider all of the available evidence on education in developing countries (see, for example, DFID's review series (The Department for International Development, 2015)). Neither do they appraise the reviews critically.

Evans and Popova's call for future work (Evans & Popova, 2015b) should perhaps be revised to encourage: "*someone to please first do an actual systematic review of real systematic reviews*". Such a work would gather the best available evidence of a topic and would highlight corroboration. But until that time we cannot accept the report's authors' claim that systematic review methods are unfit for purpose.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.envint.2016.12.020.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgements

The authors wish to thank Laurenz Langer for his contribution to the blog article upon which this article is based. We also acknowledge that the World Bank is in on-going discussions to better understand systematic review methods, and that this situation is improving.

References

Conn, K.M., 2014. Identifying Effective Education Interventions in Sub-Saharan Africa: A Meta-analysis of Rigorous Impact Evaluations.

Evans, D.K., Popova, A., 2015a. What really works to improve learning in developing countries? An analysis of divergent findings in systematic reviews. An analysis of divergent findings in systematic reviews. World Bank Policy Research Working Paper. 7203 (February 26, 2015).

Evans D, Popova A. How Systematic Is That Systematic Review? The Case of Improving Learning Outcomes: The World Bank; 2015b [Available from: http://blogs.worldbank.org/impactevaluations/how-systematic-systematic-review-case-improving-learning-outcomes.

Glewwe, P., Hanushek, E., Humpage, S., Ravina, R., 2014. School resources and educational outcomes in developing countries: a review of the literature from 1990 to 2010. In: Glewwe, P. (Ed.), Education Policy in Developing Countries. University of Chicago Press, Chicago and London.

Haddaway, N.R., Pullin, A.S., 2014. The policy role of systematic reviews: past, present and future. Springer Science Reviews 2 (1–2), 179–183.

Higgins, J.P., Green, S., 2011. Cochrane Handbook for Systematic Reviews of Interventions. John Wiley & Sons.

Hopayian, K., Mugford, M., 1999. Conflicting conclusions from two systematic reviews of epidural steroid injections for sciatica: which evidence should general practitioners heed? Br. J. Gen. Pract. 49 (438), 57–61.

Jadad, A.R., Cook, D.J., Browman, G.P., 1997. A guide to interpreting discordant systematic reviews. Can. Med. Assoc. J. 156 (10), 1411–1416.

Kremer, M., Brannen, C., Glennerster, R., 2013. The challenge of education and learning in the developing world. Science 340 (6130), 297–300.

Krishnaratne, S., White, H., Carpenter, E., 2013. Quality education for all children? What works in education in developing countries. New Delhi: international initiative for impact evaluation (3ie). Working Paper. 20.

McEwan, P.J., 2014. Improving learning in primary schools of developing countries a meta-analysis of randomized experiments. Rev. Educ. Res. 0034654314553127.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann. Intern. Med. 151 (4), 264–269.

Murnane, R.J., Ganimian, A., 2014. Improving educational outcomes in developing countries: lessons from rigorous evaluations. NBER Working Paper. w20284.

Petticrew, M., 2001. Systematic reviews from astronomy to zoology: myths and misconceptions. Br. Med. J. 322 (7278), 98.

Pullin, A.S., Stewart, G.B., 2006. Guidelines for systematic review in conservation and environmental management. Conserv. Biol. 20 (6), 1647–1656.

Steering Group of the Campbell Collaboration, 2015. Campbell Collaboration Systematic Reviews: Policies and Guidelines.

The Campbell Collaboration. The Campbell Collaboration 2016 [Available from: www.campbellcollaboration.org.

The Cochrane Collaboration. The Cochrane Collaboration 2016 [Available from: www.cochrane.org.

The Collaboration for Environmental Evidence, 2013. Guidelines for Systematic Reviews in Environmental Management.

The Collaboration for Environmental Evidence. The Collaboration for Environmental Evidence 2016 [Available from: www.environmentalevidence.org.

The Department for International Development. Collection: DFID Education Rigorous Literature Review: DFID; 2015 [Available from: https://www.gov.uk/government/collections/dfid-education-rigorous-literature-review.

The EPPI-Centre. The EPPI-Centre 2016 [Available from: eppi.ioe.ac.uk.

Thompson, R., Bandera, E., Burley, V., Cade, J., Forman, D., Freudenheim, J., et al., 2008. Reproducibility of systematic literature reviews on food, nutrition, physical activity and endometrial cancer. Public Health Nutr. 11 (10), 1006–1014.

Torgerson, C.J., 2007. The quality of systematic reviews of effectiveness in literacy learning in English: a 'tertiary' review. J. Res. Read. 30 (3), 287–315.

Woodcock, P., Pullin, A.S., Kaiser, M.J., 2014. Evaluating and improving the reliability of evidence syntheses in conservation and environmental science: a methodology. Biol. Conserv. 176, 54–62.

Please cite this article as: Haddaway, N.R., et al., "A little learning is a dangerous thing": A call for better understanding of the term 'systematic review', Environ Int (2016), http://dx.doi.org/10.1016/j.envint.2016.12.020