The Routledge Handbook of Research Methods in Applied Linguistics

Jim McKinley, Heath Rose

Systematic reviews in applied linguistics

Publication details
Ernesto Macaro
Published online on: 02 Jan 2020

How to cite: Ernesto Macaro. 02 Jan 2020, Systematic reviews in applied linguistics from: The Routledge Handbook of Research Methods in Applied Linguistics Routledge
Accessed on: 24 Oct 2023

PLEASE SCROLL DOWN FOR DOCUMENT

Full terms and conditions of use: https://www.routledgehandbooks.com/legal-notices/terms

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
What is a systematic review?

When researchers write a review of the existing research literature they do so for a number of reasons. First, they want to bring to the attention of the reader the fact that, broadly speaking, theirs is not a new topic and that it fits within extant scholarly work. Second, they may want to demonstrate that there is either a general research consensus about a particular phenomenon, or that there are conflicting findings related to that phenomenon. Third, they want to identify a gap in the research field that they believe needs filling and/or attempt to posit a hypothesis as to why there should be conflicting results surrounding the same phenomenon. Lastly, these authors will want, through their reviews of existing research, to ‘set themselves up’ for their research questions.

These types of reviews are sometimes referred to as ‘narrative reviews’ or ‘traditional reviews’. They are written by an individual author, or a small group of authors, and the reader is not told of the process that led to their writing and subsequent publication. These traditional reviews can be affected by bias and lack of systematicity: bias in terms of how many and which previously published studies are selected for inclusion; lack of systematicity with regard to whether these studies are read in any kind of depth, the extent to which they are then described in the review, how they are juxtaposed against other studies, and whether and how they are critiqued.

One of the essential ingredients of a systematic review of existing research on a particular topic is that it should be carried out with the aim of reducing as much as possible the potential bias inherent in other types of reviews. The total elimination of bias is almost certainly unattainable but a systematic review attempts to reduce bias by ensuring that a number of principles and procedures are adhered to from its very conception to the finished product.

The composition of the review team

A systematic review is almost always carried out by a team of reviewers rather than a single reviewer. Of course the size of the team will vary considerably, most notably according to the resources available, but it would be an odd systematic review that had fewer than two
reviewers in the team as one of the benefits of having a collection of minds observing a phenomenon is that they are able to apply different perspectives to it. These perspectives will have been developed through the different ways individuals have construed the relevant world. This raises the question as to who should be in a systematic review team. Let us take the case of a review of a phenomenon in applied linguistics: the acquisition of the grammar of a second language (L2). There is now general consensus in the field that the L2 grammar can either be learnt implicitly or explicitly (Ellis, 2004) although opinions vary hugely as to which process is the most effective. Now, if someone working in this field wishes to set up a team of reviewers to carry out a systematic review of research on implicit/explicit grammar learning, who will s/he invite, cajole or press-gang onto the team? In part, the answer to this question lies in the purpose of systematic reviews in the first place.

**Considering the purpose of systematic reviews**

One of the origins of systematic reviews is a concern with the reliability of research evidence (Gough, Oliver and Thomas, 2017) and particularly the way that research evidence, perhaps based on a single study, can be taken up unquestioningly by stakeholders (e.g. by doctors in the case of a study of the effectiveness of a treatment). Thus a primary purpose of systematic reviewing is to synthesize multiple studies for the benefit of stakeholders so that they can read about relevant findings in condensed form and subsequently act upon them. In the case of second language acquisition (SLA) research these stakeholders are almost invariably students, teachers and educational policy makers. It would thus make sense for a review team to include members who, at least at some point if not currently, were one or more of these stakeholders. A mixture of researchers and practitioners on a review team would not only bring to the review process ‘distributed expertise’ (Brown et al., 1993) but also greater possibility of reducing the bias that might infiltrate a team made up, say, solely of researchers. If the resources available or the situation do not permit the team to include both researchers and practitioners then it is incumbent on researchers to seek the views of stakeholders at various stages in the review process.

Of course systematic reviews are not conducted solely for the purpose of providing information and possible courses of action for teachers and policy makers; they may also be carried out to bring together the totality of the research evidence on an emerging phenomenon (for example, Content and Language Integrated Learning). This is one aspect which distinguishes between systematic reviews and meta-analyses. In a meta-analysis (see In’nami, Koizumi & Tomita, this volume) the review team examine not only methods and research findings of a series of studies but also carry out their own secondary analysis of the (almost invariably) quantitative data, in order to arrive at a ‘pooled’ data set from which meta-findings can be extrapolated. As this chapter is not concerned with meta-analyses I will not pursue this aspect further except to signal the possible bias inherent in only including studies which provide the review team with statistical data which can then be pooled. Systematic reviews can ‘pool’ the research findings but only in so far as those findings are reported by the individual studies’ authors. In that sense they have a different purpose to meta-analyses.

**The inclusion and exclusion criteria**

Prior to beginning the first search for the relevant literature, initial criteria will need to be developed in order to determine what will and will not be included. Unless there is a specific reason for doing so, most systematic reviews exclude papers which are not based on empirical
data because systematic reviews are carried out to provide a synthesis of research not of people’s opinions and theories.

One major decision that needs to be made is whether to include unpublished work such as doctoral theses. The arguments against these seem to be that a) they are very time consuming to read, and that b) they are not peer reviewed in the same way as journal articles and/or may vary in quality according to the awarding institution. In my view doctoral theses should be included as I have come across some excellent pieces of work which have not later been published in academic journals. Further, I would argue that the scrutiny brought about by examiners can sometimes be more rigorous than that of hard-pressed peer reviewers. One possible reason, sometimes articulated, for the lack of subsequent publication of doctoral theses is that the research doesn’t have definite (i.e. significant) outcomes and is therefore not worth publishing, which means that the ‘null hypothesis’ might too often go unreported. I will return to the question of what to include later when I discuss the in-depth review.

The systematic search for relevant studies

As hinted earlier, a number of factors influence how researchers of ‘traditional reviews’ or ‘narrative reviews’ go about selecting which previous studies to include in their review. Probably the most important of these is the time and resources available to them. In the second decade of the 21st century there is a vast amount of research literature available, increasingly in open access electronic format. This means that a systematic review carried out in 2018 will have very different considerations when reducing the bias of research selection than one carried out in 1998. In applied linguistics (as in many other academic fields) recent years have seen a burgeoning of new journals, some of which are available on open access, others which are not; some of which are extensively peer reviewed, others which are less so. All this available information is in addition to the personal blogs which are posted by both researchers and practitioners in the fields of SLA and/or language education as well as reports by research centres of academic institutions available online.

So how would a review team begin their search for the relevant research literature with the elimination of bias in mind? There are two basic ways in which this can be done. The first is to rely on the review team’s existing knowledge of the field and for them to agree on what the review questions will be. In other words, what will the review try to find out? This presupposes that the team has sufficient knowledge of the field to be able to formulate these questions at a level of expertise that does not exclude other questions which could be asked. The search strategy can then be devised.

An alternative and less restrictive approach is to begin with some keywords which can be inserted into reliable search engines and to see what this initial very broad search produces. In practice review teams tend to do both: have some tentative review questions in mind and then proceed to make a list of key search terms with which to test whether these review questions can be answered. The review questions are then modified as a result of the findings of this preliminary search.

One of the decisions that the review team has to make in this process is how broad the review questions can be against the constructs or the phenomena that are being explored in the research literature as it currently stands. In a recent systematic review I was involved in (Macaro, Curle, Pun, An & Dearden, 2018) we were initially trying to find out what research had been carried out in English medium instruction (EMI) in higher education. We were aware that the term ‘EMI’ was a relatively new one, although the phenomenon itself (the teaching of academic subjects through the medium of English in non-Anglophone countries) went back
many decades if not centuries. Moreover, we were aware of similar educational settings such as immersion, content-based language teaching, and content and language integrated learning, and that some research studies used the term ‘EMI’ for some of these educational settings and vice versa. In order to make the initial search as inclusive as possible we had to include terms such as ‘immersion’ as search terms.

Another consideration when wanting to ensure the validity of the systematic review (i.e. that it is reviewing what it is meant to be reviewing) is the breadth of the review questions. In principle, so as to ensure the review’s validity one would want to not make the research questions too narrow, thus missing important dimensions of the phenomenon. However, in practice, very broad research questions lead to vast collections of studies which could not possibly be given due attention. A compromise has therefore to be reached in order to make the review process manageable. The important thing to remember is to make the review process systematic and transparent. Systematicity can be achieved through extensive discussion, and indeed documentation, of procedures for carrying out the review (also known as the ‘written review protocol’). Systematicity can also be achieved by carefully documenting the decisions made as the review (and particularly the search phase of the review) progresses. In that way the review process becomes transparent to the potential users of the review or to potential replicators (or partial replicators) of the review at a later date. In other words, systematicity and transparency go hand in hand. In our systematic review of EMI we published the search terms used and as much as possible, within the word limitations, documented the search procedures we went through.

In applied linguistics the following are some of the databases that are most often searched especially if the particular field of applied linguistics of interest is SLA or language education:

- British Education Index
- Education Resources Information Center
- Language and Linguistics Behavior Abstracts
- Modern Language Association
- ProQuest Dissertations and Theses
- PsycINFO
- Social Sciences Citation Index (SSCI)

It is also a wise precaution to do some hand-searching as well by going to specific journals in order to see if, for some reason (perhaps relating to the search terms), relevant research has not been picked up by the online search.

**Dealing with the results of the first trawl**

The number of studies that are uncovered by an initial search using keywords depends on the nature of the phenomena or constructs being investigated. As already mentioned, the broader the research questions, the more keywords that review teams will have to use; the less highly defined the constructs or phenomena, the more different keywords will need to be used to get at them. Most initial searches, however, bring to the team far more studies than they can read in depth, so an initial screening has to be deployed by reading the title, the abstract and the keywords to see if the research paper is of relevance to the review questions. It is a good idea to read the title and abstract having in mind the systematic mapping activity that will follow this first or a possible second trawl (see the next section). By reading the title, abstract and keywords, the research team can then decide whether to: a) modify the review questions in order to make the review process more manageable, b) modify the review questions because as they
Currently stand they do not reflect the scope of the research that the first trawl has produced, c) modify the review questions because the terminology used in them does not appear to match the terminology being used in the field and d) modify the search terms and carry out a second search using these.

One of the decisions that will have been made by the review team is whether only one person reads the title/abstract/key words and makes a decision as to whether to keep the research paper ‘in the running’, or whether two people independently read them, make an initial decision, and then discuss those items they did not agree on – a kind of interrater reliability. Clearly the latter approach is preferable if the review team wishes to reduce the possibility of wrongly excluding an item.

Thus a lot of decisions have to be made at this stage by the review team. As always, compromises have to be made on the basis of time and resources. That is perfectly reasonable as long as these compromises are well documented and become part of a transparent review process.

**The systematic mapping activity**

As mentioned earlier, when reading through the abstracts it is a good idea to have a fairly well-developed notion of what is going into the systematic mapping of the relevant research. But what should a systematic map contain and what is its purpose? The map is basically a table which contains in very brief note form all the information needed to get an overall impression of the research. What categories to include in the map are of course dependent on the topic and the review questions but the following are nearly always found in SLA systematic maps: authors and title, date and language of publication, country/countries in which the research was conducted, the age of the learners, the educational setting (e.g. school/university), the language being studied, the type of study (intervention, relationship between variables, descriptive-only, ethnographic) the methodology used (quantitative/qualitative/mixed methods), the sample size and the claimed key findings.

If only one reviewer is entering information on the map, then, if at all possible, another reviewer should independently enter a percentage of the information and then compare. Moreover, where the information for the map is not contained in the title or abstract, the decision has to be made as to whether a reviewer should read the paper as a whole in order to get at the missing information. It would be a problematic decision to exclude from the map any study purely on the basis that it did not give sufficient information in the title and abstract.

Once the mapping activity is completed, the map serves two very important purposes. The first, clearly, is to inform the review team as to the current state of the field of research in question: how much has been done and where; the range of questions asked; the tendency to use one research methodology rather than another. This information is then used in the writing up process, usually through the use of easily understandable tables and figures, to inform future research. The second purpose of the mapping activity is to help the review team make the decision about how to conduct the in-depth review – the next phase of the review process.

**The in-depth review of existing research**

As already mentioned, a systematic online search coupled with some hand-searching of journals can uncover hundreds – possibly even thousands – of potential research reports. Even after applying the exclusion criteria that will have been established for the systematic map activity, perhaps a hundred reports will remain and will have been laboriously mapped. Most review teams will then not have the resources to read in-depth a hundred or so studies, especially when...
one applies the all important reliability criterion that each study should be read carefully and thoroughly by more than one member of the team. Whilst the systematic map needs to be comprehensive because it tells readers what has been done and what might still need to be done, the in-depth review can justifiably be more selective. Often the review questions for the in-depth review are reduced or different from the review questions posited at the beginning of the review process.

In order to reduce the number of studies, the review team will have to make very careful decisions about which of the total items that they have systematically mapped will now feature in the in-depth review. It will almost certainly be necessary to refine the inclusion/exclusion criteria and to make sure that these are reflected in the review questions. In a systematic review carried out by a team of which I was a member (Hassan et al., 2005), we explored the range of research which had been carried out on language learning strategies with a particular focus on whether strategy training (aka strategy-based instruction) was effective in improving L2 students’ linguistic knowledge and/or language skills. Thus for the systematic map we established the following inclusion criteria, which produced 38 studies.

To be included in the map, reports needed to be:

- of a strategy training intervention in language learning;
- of an intervention carried out in a formal setting such as groups of learners in schools, universities and language centres;
- a study not primarily involving bilingual learners;
- of primary, empirical research; and
- of research carried out since 1960.

However, for the in-depth review we added an additional criterion in order to ensure that the methods of the research we were reviewing were reliable: studies had to have a control or comparison group or, as it was phrased in the report, we would only include in-depth: ‘experimental studies testing the effect of the intervention against another intervention, or standard practice or no intervention’ (p. 22, my emphasis). This additional inclusion criterion reduced the 38 studies down to 25.

Having arrived at a list of studies to be reviewed in-depth, the review team will now have to decide what to look for in each study and what to extract for the purposes of reporting to the reader and for arriving at some sort of synthesis of the research and conclusions. Bearing in mind that if two members of the team will have to carry out this extraction from each study (but different pairs will be extracting from different studies) then it is absolutely essential that the team as a whole decide in advance on the categories of information that will be looked for in each study and extracted, and that collectively they will produce a template to be used during the extraction process. There will be some categories that should always be present in the document and some that will vary according to the review questions. The categories that should always be present are those relating to the quality of the piece of research itself. To illustrate this I will refer to an extraction template we used in Macaro et al. (2018) one which is adapted from a recommended one provided by the EPPI Centre (EPPI Centre, 2017).

The quality of the reporting (and possibly the quality of the research itself) can often be gleaned from the quality and content of the abstract, which should provide the reader with sufficient information to warrant further reading. Thus a comment about the quality of the abstract is important to include. Another measure of quality that needs to be considered is whether sufficient background information is given in the article as a whole. For example, does it say when the study was carried out (i.e. when the actual data was collected)? This may not seem at first glance important but some studies may be highly affected by the time period...
in which they were conducted. For example, data collected in 2001 on language learning and technology but only published in 2007 will already be out of date or at least will be read out of its historical context unless this is clearly given. Similarly, data collected prior to a major change in language education policy of a country, but published after that policy has been implemented, will need to be presented against that policy change.

Most journals now ask for a disclosure as to whether the study was funded and by whom; this is partly to make the reader aware of the possible bias that might result from the funding body putting pressure on the researchers to carry out the research in a particular way or to publish only some of the findings.

Another measure of quality that needs to be considered is whether the research is related to a recognizable theory or construct and whether this is clearly articulated in the paper. Empirical research which exists in a theoretical vacuum does not help to build up the body of knowledge and understanding necessary to advance the field of research in question. Linked to this is whether a clear and sufficient definition of the construct or constructs being investigated is provided by the study authors.

The next information to be extracted, unsurprisingly, is the topic of the study and the research questions – not always a straightforward matter, particularly in the case of the research questions. Some papers imply the research questions rather than stating them explicitly or they may limit themselves to alluding to the general ‘aims’. This is not helpful when trying to assess the relevance of the study to the questions formulated by the team for the review.

When it comes to considering the research methods used, it is a case of not just providing more information than in the systematic map, but also making some sort of judgement about the methodology adopted and whether this was adequate in: a) answering the research questions of the specific study under consideration and b) whether the methods were relevant for answering the questions that the review as a whole is asking. For example if a study was asking how effective a particular type of teaching activity is for learners to acquire vocabulary implicitly, they may use a methodology whereby learners are exposed to large quantities of L2 text and, using pre- post- and delayed-tests of vocabulary, they can determine (whilst controlling for possible confounding variables) how much and what type of vocabulary can be learnt simply through extensive reading. This would be a perfectly reasonable methodology to use for the research questions of that specific study. However, if the systematic review questions asked whether L2 vocabulary is better learnt and retained either by a teacher providing L2/L1 vocabulary lists (and strategies for learning that vocabulary) or only providing students with multiple texts in which that vocabulary was embedded, then the individual study’s methodology would be only partly relevant to the aims or the systematic review. This does not mean that the methodology used by the individual study was defective, merely that it was asking a somewhat different question to that being asked by the review team. Thus whilst the study might be included in the in-depth review its limitations in answering the review questions would be highlighted.

Another aspect of methodology that a review needs to extract is the sampling strategy of a particular study. This is of special importance in reviews of SLA, which often try to arrive at an international picture of a particular L2 teaching and/or learning phenomenon. The reader of the review will want to know something about the population being studied in a particular study. Who are these people? What is the context in which they are teaching or learning? How is that context different from other populations in other studies in the review? Moreover, the quality of a particular study can in part be judged by the relationship between the population and the sample, through the sampling frame adopted. This will allow both the authors of the study and the review team to make a judgement about how generalizable the findings of the individual study are and whether they are generalizable (through the systematic review) across different L2 contexts.
Having carefully considered the methodology used by the researchers of the individual study, the review team will need to consider the actual findings and whether these appear to match the methodology used – in particular whether the relationship between the findings and the conclusions, as drawn by the authors, is the same as that drawn by the review team. I have often come across conclusions drawn which to me seem not to match the results of the investigation, or seem tangential to them. Particularly I recommend taking a look at the kinds of words used in what I call the ‘final moves’ (the final concluding statements made by the authors) to see whether any personal bias appears to be creeping in. We are all capable of this personal bias and that is one of the great advantages of having a review team looking at the results and conclusions rather than a single individual.

Writing up the in-depth analysis of the review

There are three possible approaches to how the in-depth analysis should be presented to the reader once all the extraction templates have been completed and discussed. The first is to group the studies under categories or under the review questions and then to report them one by one with varying degrees of critical analysis. This might seem a very boring way to go about it, but it may help the reader-researcher who wants to scan the field in order to begin thinking about his/her own research.

A second way is to further group the studies according to the research findings – so that they aggregate the evidence. The grouping may show that there is a divergence of findings, in which case the review team has to try to work out why that divergence is occurring (a population-based divergence or a research-methods based one, for example).

The third way is to group the studies under the methodologies used. This would be an approach that might be appropriate if, for example, the review team’s aim was to explore a relatively new phenomenon or construct, or a construct which they believed had been studied and documented without giving due attention to establishing what the construct actually was in the first place. For example, there might be two sets of previous research. The first investigates the construct only through the use of quantitative research methods (e.g. language tests) and comes up with one set of findings. The second investigates the construct through ‘learner-internal’ qualitative research methods (such as interviews and stimulated recall) and provides insights which contradict the quantitative findings or at least give much greater understanding of the possible variations in the findings.

One constant in systematic reviews is that there should be a neat and accessible summary of the review, one written directly with stakeholders in mind who may not have the time, the desire or the expertise to read the whole review. This section sometimes is placed at the beginning of a review and is called the ‘executive summary’. This is usually found in reports published by organizations such as the British Council or the World Bank. In these executive summaries the rationale for the review and the review methodology are provided in non-technical condensed form. In the case of systematic reviews published in academic journals (and note that these are in the minority), word count limitations preclude the use of an executive summary. Thus for these the ‘summary of the findings’ is best placed at the end before the discussion section and the conclusion.

The discussion and conclusion sections of a systematic review will be different from those of individual studies. In the latter the authors are expected to match their findings to previous research. In the case of a systematic review, the matching will already have been done in the main reporting sections. Unless there have been prior systematic reviews (in which case some comparison of review findings can be made), it remains for the review authors to stand back
and comment on the field as a whole, to synthesize the totality of the findings and to identify what we might call ‘macro-gaps’ rather than ‘micro-gaps’. For example, in our review of EMI we set the kind of research done in higher education against a backdrop of the research done in secondary education and found some notable but not necessarily justified differences, particularly with regard to neglect of certain geographical regions.

The other requirement of the discussion and conclusion sections is a statement regarding the contribution and the quality of the research in that particular field and how it could be improved. The decision that the review team will have to make is whether they judge, rank and also name each individual study, or whether they provide a summary of the overall contribution and quality without naming each individual study. Before commenting on that decision, let us first consider how to go about making a judgement on contribution and quality. Again adapting from the EPPI Centre guidelines, what follows are the four criteria that can be used in order to judge the quality of individual studies on a scale of high/medium/low. The first is the level of relevance an individual study has to the systematic review. One way to consider this is the extent to which the research questions of the study match one or more of the review questions. Another (slightly more humorous) way is to think about how excited the reviewer got when s/he first came across the title/abstract/key words of the individual study. S/he may have subsequently been disappointed on reading the whole article but at ‘first glance’ it looked as if it was highly relevant for the review. The disappointment may not be because of the quality of the research carried out – merely that it was not as relevant to the review as was at first thought.

The second category, linked to the first, is whether the methodology used in the individual study was the kind that would be appropriate in answering the review questions. Clearly a study that asked students what they believed would be a better teaching method would not be appropriate in informing a review that asked about the effectiveness of interventions in SLA. Again, that does not mean that the quality of the research design of the individual study was at fault, merely that it did not match what the review was looking for.

The next criterion is the really tricky one: taking into consideration all aspects of the individual study’s methodology, can the study findings be trusted in answering the study question(s)? This is essentially an evaluation of the overall quality of the individual study. Here the researchers involved in the individual study are being evaluated by the review team. This raises two important issues. First, in some systematic reviews the overall quality of the individual study is used as an inclusion/exclusion criterion. This is problematic as it means that some possibly valuable information (for example about a particular population) is being denied the reader on the basis of a quality judgement. The second issue is whether the individual study should not only be graded but also named. In Hassan et al. (2005) we did both grade and name; In Macaro et al. (2011) and Macaro et al. (2018) we only graded. I would now opt for not naming, for three reasons. First, grading a publication as ‘low’ (i.e. not trustworthy) seems to me to be not only offensive to the authors of that paper but also arrogant on the part of the reviewers, who are, after all, only making judgements as mere mortals themselves. Second, given that the reader of the review is being provided with an overall grading of studies according to four criteria, then the aims of the systematic review have been fulfilled: the reader knows where the strengths and weaknesses are. Third, in the narrative of the systematic review some critical analysis of each study should in any case be provided (e.g. ‘note that this intervention study did not use a control group’). Thus if readers wish to follow up the critical analysis in greater depth they can do so by reading the article themselves.

The last criterion in the judgement of contribution and quality is a summarizing one. It asks: is the contribution of the study to answering the question/s of this specific systematic
review, high, medium or low? Note that this judgment is not necessarily evaluating the individual study for its internal quality, merely its ‘external’ contribution to the review.

Conclusion

A systematic review is carried out both for the benefits of other researchers and for stakeholders who may not be involved in research but are users of the research evidence in their field of work. In that sense a systematic review should not only be objective, reliable and transparent in the message it sends out but also should project a sense of accountability. The review team is providing a service to the end user, going beyond the individual study or even groups of studies with their limited scope and contexts. This synthesis brings with it a responsibility. Lack of responsibility and accountability is no less dangerous in the field of SLA or language education than it is in the field of medicine. Busy medical practitioners do not have the time to sift through dozens of studies researching a particular illness and potential cure; they rely on systematic reviews.

Like individual studies, a systematic review should declare the limitations of the review. Some of these can be beyond the control of the review team (e.g. inaccessibility of material written in a language in which they do not have expertise). However the quality of reviews themselves can be mediated by high standards of project management: careful planning, record keeping and the avoidance of ‘corner-cutting’.

Systematic reviews of applied linguistics – if carried out responsibly and transparently – can be of great benefit to related professions. They can avoid precipitous policy implementations which can occur subsequent to results of individual studies and ultimately can lead to narrowing the credibility gap between research and practice.

Notes

1 An example of this can be found at (EMI Oxford) group web page.
2 This is available at (EMI Oxford group web page).

References


