To save the research literature, let's make literature reviews reproducible

*Last week the Impact Blog featured a post from Richard P. Phelps, in which he proposed that journals [get rid of their requirement for a literature review](). **Arnaud Vaganay** agrees with much of what Phelps said, literature reviews are erratic and self-serving, but suggests doing away with them altogether is likely to make science less efficient and less credible. Instead, literature reviews and their component parts should be thought of as reproducible research assets, just like the protocol and dataset used in the empirical part of a study. Cumulative literature reviews should be based on a methodology that is more systematic, reproducible, and reusable than the traditional, narrative literature review.*

In a recent Impact Blog [post](), Richard P. Phelps proposed that journals drop their literature review (LR) requirements. While I agree that most LRs bring little value, I am afraid that, given the cumulative and replicative nature of science, such a drastic measure would do more damage than good. When done well, LRs can be very useful to researchers, readers, and meta-researchers. A more principled approach is now available.

There *is* something wrong with LRs

Let me start by saying that I agree with many of the points Phelps makes. First, I agree that most LRs lack rigour. A LR is an analysis of the existing literature on a given subject. Like any analysis, it is based on three components:

- *Data*: individual studies.
- *Methods*: decision criteria used to determine whether or not a given study should be included in the LR, how methodologies should be compared, etc.
- *Results*: a factual report of what makes the new study *replicative* and what makes it *innovative* (bearing in mind that the ideal study is a mix of both).

Like any other analysis, the LR should be transparent and reproducible. Unfortunately, it rarely is. I do not recall ever reading a LR that includes the search string and the list of bibliographical databases used by the authors. [I am not the only one](). Of course, I am not talking about systematic reviews, although they are [not always as reproducible as they claim]().

Second, I agree that claims such as "this is the first study" or "no previous studies" are often unhelpful. The problem is that they are unfalsifiable, disingenuous, and ultimately misguided. They are unfalsifiable, because the data and the methodology are almost never shared. As a result, there is no way the reader can challenge the authors' assertions. They are disingenuous, because they are based on a narrative rather than a protocol. Any inconvenient result tends to be swept under the rug. Lastly, they are misguided, because science needs replication and innovation in equal parts. A study that is 100% replicative is a waste of money. Conversely, a study that is 100% innovative (assuming for a second that this is possible, which it is not) can neither be interpreted nor appraised. Any study is a mix of old and new, and an honest LR is one that clearly indicates what components were created and what were reused.

Third, I agree that there is insufficient quality control. Several studies have shown that studies with positive results were significantly more likely to be cited than similar studies with neutral or negative results (a systematic review can be found [here]()). This suggests that editors and peer reviewers do little to ensure that LRs tell the whole story rather than a good story. It is hard to blame them; by some accounts the [scientific output doubles every nine years](). As it becomes increasingly difficult to stay on top of the literature – even in one's own field – editors and peer reviewers should pay more attention to the way LRs are conducted.

LRs are our friends

LRs are more useful than many researchers realise. First, they make research incredibly more efficient. They can be used to demonstrate what samples were used in previous studies, what assumptions were made and/or tested, how variables were measured, what analytical methods were used, what results were found, and what mistakes should be avoided. Ignore this at your peril.

Second, they are essential to build trust and confidence, which is what all science is about. As an author, I am much more relaxed when I can prove that a given method has stood the test of time. I am also more confident when I claim that such or such approach has, to the best of my knowledge, not been attempted before.

Third, LRs are essential to interpret empirical results. I believe that "no isolated experiment, however significant in itself, can suffice for the experimental demonstration of any […] phenomenon" (words attributed to Ronald Fisher by Morrison and Henkel). The "one chance in a million" will undoubtedly occur, which is why researchers have the moral duty to question new results (scepticism). It is only possible to discuss the credibility of new results through systematic comparisons with previous results. The LR is the part of the manuscript that identifies these results.

Cumulative LRs to the rescue

Readers who are as frustrated with the current state of LRs as Richard Phelps and me might be interested in Cumulative Literature Reviews (CLRs) (see from slide seven onwards). CLRs start from the premise that LRs and their components (literature dataset, LR protocol) are *research assets*, just like the protocol and dataset used in the empirical part of a study. They are assets in that they can be reused – and *should be* reused – by colleagues sharing your research interests. Think about the time you could save if all you had to do was to update a LR rather than do it from scratch (like dozens of researchers before you)! For those who agree with this vision, CLRs provide a 10-step methodology that is more systematic, reproducible, reusable, factual, and visual than the traditional, narrative LR. Importantly, this method can also be taught, delegated, and improved. The methodology is still at the development stage and is being piloted in two ongoing studies. Questions and comments are more than welcome.

Let's not throw the baby with the bath water

I agree that most LRs are erratic and self-serving. But getting rid of them is likely to make science less efficient and less credible. If we agree that LRs can be helpful and can be improved, then all we need to do is raise our game.

Note: This article gives the views of the author, and not the position of the LSE Impact Blog, nor of the London School of Economics. Please review our comments policy if you have any concerns on posting a comment below.

Featured image credit: Alfons Morales, via Unsplash (licensed under a CC0 1.0 license).

About the author

Arnaud Vaganay *is a methodologist and meta-researcher. He is the founder and Director of Meta-Lab and a Catalyst of the Berkeley Initiative for Transparency in the Social Sciences (BITSS).*