

Room For Writing - Approaching a literature review

Approaching a literature review

Thanks to Professor Patrick Dunleavy for the LSE for this piece, excerpted from a longer article published as a blog post on medium.com

The original blog was aimed at specialist researchers and PhD students, but the perspective is relevant to those doing research at whatever stage of their academic career - from first year undergraduates onwards.

It's quite a long read (the original blog was rated at 13 minutes), but its worth the effort - and you may find it reassuringly closer to your own existing practices than some views :-)

Extract from Doing faster literature reviews

Instead of freezing our understanding of a field at one time, often indeed a time when we least understand the field, we should see the literature review as a repeated component of any ongoing research. We need more agile ways to surface other relevant research at every stage of our thinking and 'writing up', not just at the outset. We also need to consider how researchers actually work now, which is not very well presented by most institutional advice webpages or courses, generally produced by librarians I think, rather than by creative researchers themselves. So, in hopes that it will trigger some pushback comments and reactions, I set out here a first (deliberately controversial) attempt to outline strategies that contrast with the rather orthodox (and perfectionist) advice that seems to be out there at present.

1. Use Google search tools first and foremost. This may seem controversial to most librarians, who want researchers to use the proprietary bibliometric databases that they have expensively acquired, and sometimes researched about. But Google tools are clearly the best available in many dimensions and most disciplines, and they are easy to use in common ways, universally available on any internet PC or tablet, and free.

Google Scholar is by far and away the best (most inclusive) of the world's bibliometric systems now, chiefly because it covers not just journal articles, but also citations of books and the many 'grey literature' reports originating from academic sources. You can search datewise in GS and it is usually fair enough to date restrict research searches to the last five or six years. GS also responds well to Boolean algebra search terms, putting linked terms in "double quotes", and using operators such as AND, OR and NOT. You need to try a wide range of permutations of possible search terms, and to refine the combinations looked for in line with what the GS results are throwing up.

Once you have well-defined search combinations, a realistic goal then is to make sure that you skim through the first 200 or 300 results. (If you set up your page to show 10 search results in snippet mode at a time then this is only 20 keystrokes). Typically there are likely to be three levels of results. (i) 'Remote possible' papers where you can just paste the GS 'snippet view' details into your 'Materials' file or archive, in case they might later prove interesting; (ii) Partly relevant papers, where you might copy across the article title and abstract only; and (iii) Clearly relevant papers, where you download the full text to your PDF library.

Google Books is an essential additional tool in the humanities and social science. Even in STEM disciplines it can be a useful add-on resource when seeking textbooks (best for explaining new materials), or the occasional 'summation', think piece or research commentary books from senior scientists.

Essentially Google has now run around 10 million books through optical character readers so as to create online images of each page. For books that are out of copyright, Google makes available the full text for reading online, but the material cannot be downloaded in the free use version of the program. The text of most out of copyright books is also fully searchable, so you can easily find specific sentences, quotations, or words of interest anywhere in the book.

For books still in copyright how much information is viewable on Google Books depends on what agreement the book's publisher has reached with them.

(i) The most restrictive 'no preview' option just replicates the publishers' book blurb and perhaps gives the contents pages.

(ii) The 'snippet view' offers only a few short glimpses of the book's content. But this still allows readers to word search the full text for terms or phrases, and so assess how much coverage there is of relevant material.

(iii) The most expansive Google Books preview shows many full pages of the text, but leaves out some

Room For Writing - Approaching a literature review

key chapters or sections. However, you can still use the word search across these omitted sections, and get a snippet idea of what's covered outside the full text pages.

In either the snippet view or full text preview you can't copy any text from Google Books. But in Windows, press Control+Print Screen to capture your screen view and then copy that image into a Word archive file. If a book covers what you are interested in only briefly, it's easy to copy across these few pages of relevant text as a succession of images, obviating the need to consult the text itself. The text-finding software in Google Books is so powerful that many scholars now use it as an online index to find material within books already on their shelves, but which have either no index or the normally very inadequate academic book index system.

The main Google (Web) database has some specific search advantages compared to Scholar and Books, notably in being much more up to date, and in covering news media, blogposts, and the extensive 'grey literature' from corporate and professional bodies (as well as academic reports and working papers covered with a lag in GS). Because of its enormous size, however, main Google is best searched with relatively extensive and if possible distinctive phrases. For instance, searching for the author- and concept-distinctive jargon term 'deliverology' would be feasible, where the simpler and more ordinary language term 'delivery' would generate far too many entries to be useful. If your search initially throws up thousands of items, add additional content-distinctive words (again using Boolean operators) to get the numbers down to a feasible size.

Similarly if you cannot find a quoted author or source, but you do have a quotation of at least five or six words, search for them inside "double quotes" and you'll be surprised how quickly this thins down the lists of Google's possibles.

Google Scholar Citations is a very helpful search extender. It's a database of authors that most leading researchers now have an entry on, and which Google auto-updates so that it is always current. (If you're a PhDer or researcher who's not yet joined GSC, are you perhaps an academic hermit? Don't be).

Once you've got a list of 'core' articles or books directly relevant to your research from the sources discussed above, look up all the key authors on GS Citations to see if they have other publications on the same theme. GSC also shows you how many citations a given source has, so you can see roughly how important it may be. Sources that are heavily cited (given their age) generally deserve more care and attention. Often a major author may also have several versions of their argument, where generally the older one is more cited (because recent work takes quite a time to acquire citations).

Outside STEM subjects key authors may well have both a book and article versions of the same work. Of course, different discipline groups also have varying citation rates, that you need to control for. Yet used intelligently citations levels help you cut through the problem of distinguishing remotely what is or is not likely to be important. GSC also lists all the other works in Google Scholar that have cited X's key work (click on the cites number to see a list of these). For work directly related to your research, these listings have a high probability of including other relevant stuff, so it pays to search them fully and thoroughly. GSC also lists co-authors—so if you find that X has written well on your topic with Y, look up Y's publications as well on GSC. For journal articles GSC has a powerful Google Scholar Metrics tool. Just enter the name of a journal you are unfamiliar with, and GSC will show you excellent indicators of its importance in its home field.

Finally, the GSAlerts service provides excellent personalized updates to researchers in line with their publications, and with the kinds of authors they are following. Especially if you have published a lot (so that GS has a lot of information about your research interests to go on) these Alerts are phenomenally accurate and time-saving. They may not be so effective for new researchers though. Taken together, these tools help you do a digital equivalent of 'searching along the shelf' in the library, but in a much faster way and with far more useful contextual information.

2. Learning to use proprietary databases, such as the cross-disciplinary World of Science and Scopus, and more single-discipline or topic-focused resources, such as PubMed, can take time. These expensive and charged for databases are all human-compiled and so require that your university library has a subscription for you to gain access. They generally focus on articles in academic journals judged 'reputable' on conservative criteria. It can be time-saving and helpful to know that what elite Western academics might judge as 'marginal' work has mostly been excluded, but these databases still include a range of materials sufficiently wide to make it vital that you judge directly the quality of what

Room For Writing - Approaching a literature review

you find.

The biggest advantage of the proprietary databases is including new literature from core journals quickly, an especially important feature in STEM disciplines. Yet their 'legacy' designs often predated the modern digital era, making them 'clunky', quite difficult to use, and different from one to another. Because they are hard get familiar with initially, and re-familiarize with after a break, you normally have to be trained in the Library about how to use them creatively. So these systems tend to reinforce the idea of a literature review as a discrete phase of research.

Outside STEM sciences, the coverage of many edited databases like these is quite poor, with low 'internal inclusion' levels (an index showing how many references from articles included the database are also sources found within it). Not covering books and chapters in books is a big problem in most social sciences and all humanities disciplines. A literature review compiled on this basis might be worse than useless, because it is actively misleading. Scopus has included some books for some time, and Web of Science has been trying to reverse this blank spot recently. But neither comes close to the comprehensive coverage in Google products.

Similarly, if your research focuses on regions outside Europe, north America or Australasia, it is important to recognize that despite their size, the main proprietary databases cover only a small proportion of published work about these areas—perhaps as little as 3 per cent of all journal articles worldwide. They are also English-biased sources, which may matter a lot if a lot of research in your topic area is published in other languages.

3. Open access sources matter increasingly, chiefly because they give you quick access to full texts. They range now from 'born open access' journals (like plusone, now the biggest journal in the life sciences area); through web-based alerting and discussion paper series (like Arxiv in physics and neighbouring areas); over to leading blogs (such as EUROPP, a leading European blog curated by LSE).

Archive sites that store full text versions of articles, chapters and papers are also very important, with ResearchGate running an especially good service that alerts you when another researcher you are following deposits any new materials. Twitter and Facebook streams linked to blogs, or based around other open access databases, also now serve as very important means of academic communication in the social sciences and humanities. Many older researchers have bewailed the increasing volume of academic work and their inability to keep pace with it. But the full range of search extension services should actually mean that you can keep up to date more easily across a far wider range of materials than in the past.

4. Store full texts digitally (and don't make conventional notes for recall reasons). At research level, details and precision matter—for instance, how you design an experimental protocol; how strong different observed effects are; or how to detect and interpret complex multi-causation patterns. It's no good making conventional notes about such things. If a source is directly relevant for your research, you need to store it securely on your own Cloud space or hard disc drive; you need to see exactly what the original author(s) said; you should comment directly on it so as not lose your reactions and questions; and you need to have both text and comments constantly available to you, to consult whenever you are citing or covering that source.

Modern bibliographic management tools also go far beyond the old confines of reference-only systems like EndNote and RefWorks. The package Zotero helps you clip, annotate and store all forms of material that you read digitally, annotate it and then manage your accumulating collection. Not everyone gets on with it though.

Mendeley lets you store PDFs and other documents in along with details that are ready-formatted for re-expressing in bibliographies of different formats. You can also of course store all your own publications with them, and create groups that have collective resources. Each of these packages takes time and often training to understand, but for well organized people they should repay early learning and input costs with improved later access to sources.

Patrick Dunleavy, 2014 ¹

¹ The medium, doing quicker literature reviews <https://medium.com/advice-and-help-in-authoring-a-phd-or-non-fiction/doing-a-quick-literature-review-f94700f947ce#wgmjn43of> December 8, 2014