# Literature review papers
# the search and selection process

van Wee, Bert; Banister, David

**DOI**
10.1080/12460125.2023.2197703

**Publication date**
2023

**Document Version**
Final published version

**Published in**
Journal of Decision Systems

**Citation (APA)**
van Wee, B., & Banister, D. (2023). Literature review papers: the search and selection process. *Journal of Decision Systems*, *33*(4), 559-565. https://doi.org/10.1080/12460125.2023.2197703

Taylor & Francis
Taylor & Francis Group

Journal of Decision Systems

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tjds20

# Literature review papers: the search and selection process

Bert van Wee & David Banister

**To cite this article:** Bert van Wee & David Banister (2023): Literature review papers: the search and selection process, Journal of Decision Systems, DOI: 10.1080/12460125.2023.2197703

**To link to this article:** https://doi.org/10.1080/12460125.2023.2197703



Published online: 11 Apr 2023.

JOURNAL OF DECISION SYSTEMS
https://doi.org/10.1080/12460125.2023.2197703

Taylor & Francis
Taylor & Francis Group



# Literature review papers: the search and selection process

Bert van Wee [a] and David Banister [b]

[a]Transport Policy, Delft University of Technology, Delft, Netherlands; [b]Emeritus Professor of Transport Studies, School of Geography and the Environment, University of Oxford and Senior Research Fellow at St Anne's College, Oxford, UK

**ABSTRACT**

In an earlier publication (Van Wee & Banister, 2016), we explained that literature review papers (LRP) should explicitly report on the search and selection process of documents included. In this paper, we present two approaches for reporting the methodology for doing this: a basic approach and a reasoned approach. The basic approach includes reporting databases(s), keywords, search strings, snowballing, the selection of documents, presenting an overview of documents included and reporting additional selection criteria (if applicable). The reasoned approach adds to the basic approach (what was done) by explicitly explaining the motivation for choices and showing the selection process graphically (why it was done). The two approaches should be seen as options, and not the only alternatives. We recommend that authors of LRPs depart from these approaches and modify them where appropriate. The important lesson is that authors should be explicit about the approach adopted, as this enables the reader to understand the thinking behind the LRP and the conclusions drawn.

**ARTICLE HISTORY**
Received 10 October 2022
Accepted 28 March 2023

**KEYWORDS**
Literature review; methodology; basic and reasoned approaches

## 1. Introduction

Literature review papers (LRPs) form a major contribution to the output of scientific research. They introduce readers efficiently into a specific research area by providing an overview of the state of knowledge. Literature reviews should not only provide an overview but also add value, examples being the gaps in research and a research agenda, an evaluation of the methods applied or the theoretical underpinnings, or the implications for practice (see Van Wee & Banister, 2016, for options to add value in a literature review paper).

LRPs are now an essential part of the research process as they provide an accessible, up-to-date and relevant introduction to a research topic. This provides an invaluable starting point for young researchers looking for inspiration and information on a topic, but LRPs are equally important for more seasoned researchers seeking to move to a 'new' topic or for researchers returning to a topic after a period of time. Many of the most interesting research questions lie at the interface between different disciplines and literatures, and there is now so much 'knowledge' that is

**CONTACT** Bert van Wee g.p.vanwee@tudelft.nl Transport Policy, Delft University of Technology, the Netherlands

available across many different formats. A high-quality LRP allows access to a topic through the perspective of an established expert, and this in turn means that much of the necessary groundwork has already been carried out, thus making it easier to find a way into that topic. The LRP should highlight opportunities and potentially useful lines of enquiry (the research agenda), but it can also identify potential cul-de-sacs. LRPs form an essential foundation and reflection on research topics, and they should be seen as a valuable part of the research process.

Until recently, it was quite common to not be very transparent about the methodology of an LRP. This is surprising because there is a quite strong agreement that scientific research should be reproducible. In 2016, we wrote a paper to provide authors of LRPs guidance on how to do the research for an LRP, and how to write their paper (Van Wee & Banister, 2016). Since then (to March 2023), this paper has been downloaded over 480,000 times (almost eight times more than the second most downloaded paper in Transport Reviews) and cited over 250 times (in SCOPUS), and we have the impression that it has proven to be helpful to people who want to write a literature review paper.

In that paper, we did not explicitly provide guidance on how to report on the methodology used in a condensed and structured form. In this paper, we propose two approaches for reporting the methodology, labelled 'Basic Approach' and the 'Reasoned Approach'.

Our expertise is in transport. The way LRPs are presented differs between disciplines. LRPs in the area of transport are similar to many other areas of social science, such as geography, psychology, innovation sciences, economics, and environmental sciences, and to the best of our knowledge also in areas like information systems, decision support, and management. However, LRPs in other areas, for example, health science, often differ in these social science areas. We aim to at least serve the areas in which LRPs are quite similar to those in transport. Section 2 presents the two approaches, and Section 3 comments on the issues raised.

## 2. Two approaches for reporting the methodology of LRPs

These two approaches use the same thinking, but the Reasoned Approach gives more detail (Table 2) about why certain pathways were followed, whilst the Basic Approach (Table 1) just reports on the six elements covered. These elements are common to both approaches, namely databases, keywords and search strings, snowballing, selection, results, and additional criteria. For additional information on searching and selecting sources we refer, for example, to Moher et al. (2009), or Van Wee and Banister (2016).

The three most important topics, namely the selection of databases, the search strings, and the selection criteria of publications found are now discussed in more detail. Databases in which one searches for documents should always be made explicit in an LRP. The most common databases are Web of Science (WoS – www.webofscience.com), SCOPUS (www.scopus.com), and Google Scholar (https://scho

**Table 1.** The basic approach for reporting the methodology of a literature review paper.

| Elements | Basic approach | Example |
|---|---|---|
| Databases | List the database(s) used for the search of papers | SCOPUS and Web of Science |
| Keywords and search string(s) | List the keywords and search string(s) | (transport*) AND (health OR exercise) |
| Snowballing | State if you applied forward and/or backward snowballing | We applied forward and backward snowballing departing from the references listed in Table [table number]. |
| Selection of documents | State the principles on which you selected documents | We selected papers based on titles, keywords and abstracts. |
| Results: documents found | Report the documents selected | A table reporting author(s), year of publication, topic, geographical scope, method(s) used |
| Additional selection criteria (if applicable) | Make explicit. For example, language, time frame, geographical scope | We include English language documents only, published since 1990 related to the topic in OECD countries |

**Table 2.** The reasoned approach for reporting the methodology of a literature review paper.

| Elements | Reasoned approach | Example |
|---|---|---|
| Databases | List the database(s) used for the search of papers | SCOPUS and Google Scholar |
| | Explain your choice | We first searched documents in SCOPUS. Because the topic of this LRP, Artificial Intelligence, is rapidly evolving and academic publications might miss recent trends, we additionally searched for papers published in the past 24 months via Google Scholar. |
| Keywords and search string(s) | List the keywords and search string(s) | (transport*) AND (health OR exercise) |
| | Explicitly discuss the motivation. For example, discuss synonyms, different terms in different research areas that have about the same meaning. | Because not all relevant papers use the term 'accessibility' we also searched for literature using related terms, i.e. 'connectivity' 'access' and 'proximity'. |
| Snowballing | State if you applied forward and/or backward snowballing | We applied backward snowballing departing from the references listed in Table [table number]. |
| | Explain your choice | Because the first paper in this area were published less than 6 months before the date of searching, we only applied backward snowballing. |
| Selection of documents | State the principles on which you selected documents | We selected papers based on titles, keywords and abstracts. |
| | Include a table showing, or figure visualising how many papers were added/removed after each selection step | See Le et al. (2022) (Figure 1) and Table 1 |
| Results: documents found | Report the documents selected | A table reporting author(s), year of publication, topic, geographical scope, method(s) used |
| Additional selection criteria (if applicable) | Make explicit. For example, language, time frame, geographical scope | We include English language documents only, published since 1990 related to the topic in OECD countries |
| | Explain your motivation | We only included documents published since 1990 because the methodologies used before that data are outdated – this would need to be supported by a key reference |

lar.google.com/). WoS has more older papers than SCOPUS, and SCOPUS is a bit more inclusive SCOPUS also includes some other sources in addition to papers in ISI journals, such as books published by academic publishers. Google Scholar also includes 'grey literature' (defined by Wikipedia, assessed 13 September 2022, as 'materials and research produced by organisations outside of the traditional commercial or academic publishing and distribution channels'). For comparisons between databases: see, for example, Singh et al. (2021) or Martín-Martín et al. (2021).

The academic literature often lags behind in some respects, first because authors often publish a first version of their work at a conference or in a report, and later upgrade their work and submit it to a journal. Then, the whole process of reviewing, revising, and publishing can easily take 1 year or longer. Consequently, very recent work can be traced better via Google Scholar than in SCOPUS or WoS. For rapidly emerging topics, such as (at the time of writing this paper, 2022) Artificial Intelligence, Machine Learning, or Mobility as a Service, it could be an option to search not only in SCOPUS and/or WoS but also in Google Scholar. The use of Google Scholar is also recommended for non-academic publications, such as policy documents.

With respect to the search strings, keywords need to be reported. Do not use formulations like: 'we used keywords like [list]', or formulations like 'we used as keywords [A, B, ... } and synonyms'. The reader needs to know which synonyms. Preferably copy-paste the search string that you used exactly, for reasons of reproducibility. For the selection of paper found via search strings, in the Reasoned Approach we suggest that a figure is used visualising how many papers were added/removed after each step in the selection process. It is important to make all search strings explicit, partly so that the analysis is reproducible, and partly to reflect on the importance of the choice of the keywords and search strings selected. In some cases, these might be relatively straightforward, but in others testing alternatives might be fruitful to give confidence in the search strings used. In addition, limitations relating to language, the time frame used and the geographical scope need to be imposed, both to restrict the search process and to acknowledge the boundaries of the authors' competences. In all cases total transparency is essential (Tables 1 and 2).

As an example of the visualisation of the selection process, a recent study by Le et al. (2022) is used – see Figure 1. These three authors were investigating the rise in e-commerce and the impact on personal travel behaviour, and as can be seen, they systematically identified papers through database searches and other sources. This gave them more than 2000 possible papers to review. They first reduced this through a systematic screening of titles and then abstracts, and then further narrowed their search by reviewing the full texts of about 100 papers. Finally, they decided to include 42 papers in their LRP derived from their selection process. This was supplemented by a small number of additional papers identified through a parallel snowballing exercise. The logic of the whole process is transparent and reproducible. In many cases, a table making the selection process explicit in several steps (rows in the table) can also do the job.

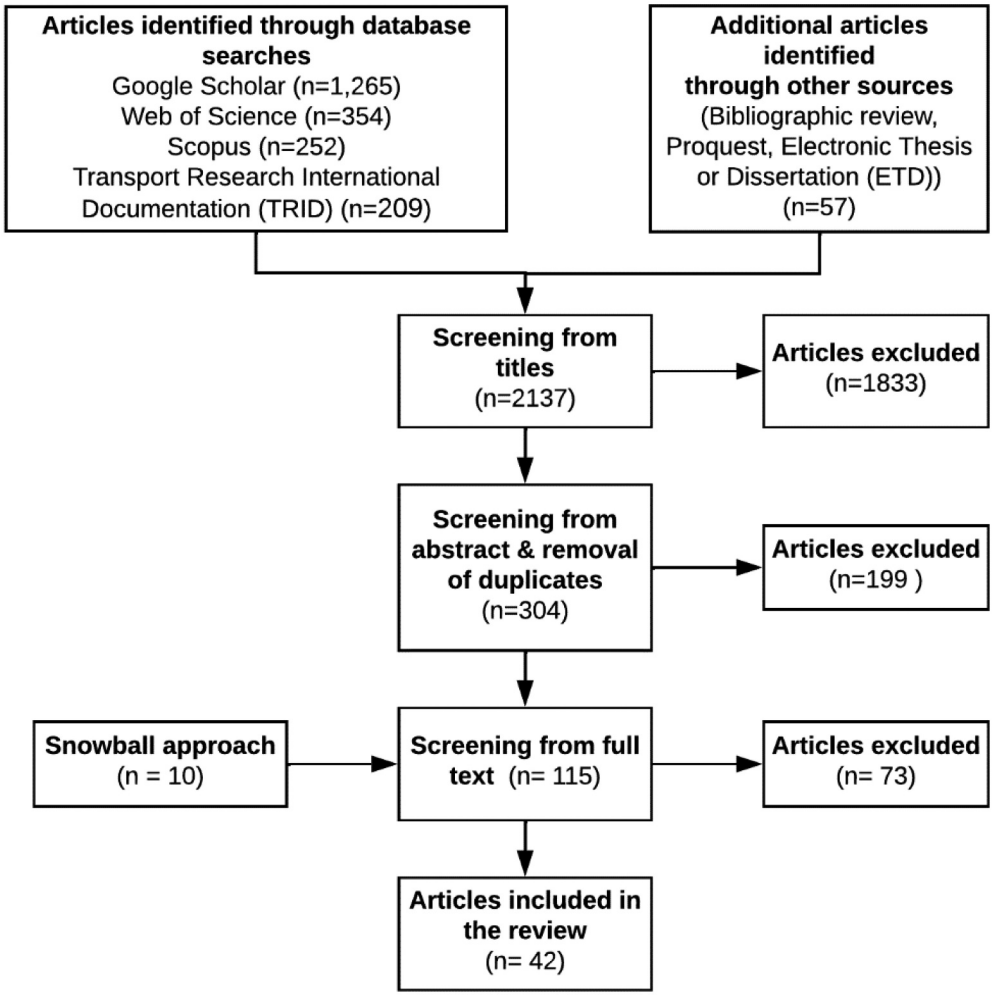


Figure 1. an example of the visualisation of the selection process (Le et al., 2022).

## 3. Conclusions

Our main message is that we recommend authors to be very specific with respect to reporting the methodology of an LRP, especially the process of selecting documents. The full methodology is not only limited to document selection, but it can also include several other research steps like the categorisation of documents (e.g. by period and region), the interpretation of qualitative and/or quantitative findings, or the coding of findings. These methodological steps are also very important and should be reported explicitly, but the heterogeneity in other methodological steps involved in searching for and selecting literature is very large, and it is beyond the scope of this paper to suggest possible alternatives.

Note that the two alternatives we suggest should not be treated as the only two approaches, more options are possible, and the approaches we provide should be treated as a point of departure, not as a prescriptive blueprint. Authors of LRPs can depart from

one of these approaches and modify them to suit their own requirements. For example, they could report explicitly how they clustered all selected documents in categories, plus why.

If researchers have specified their research steps, they could ask someone else to repeat the search process and the selection of documents based on the methodology reported. Ideally, that person would come to about the same selection, and if this was the case, it would be a nice validation of the methodology used. But it could also be a time-consuming step. We recommend authors of LRPs to debate our two proposed approaches. Based on their experiences, updated versions could prove to be even more helpful in the future.

We expect automated searching, selecting and screening to become more important in the future. For example, Chai et al. (2021) discuss the screening of abstracts making use of Machine Learning. At the time of writing this paper (2022), automated searching, selecting and screening have not been a generally accepted and used method, at least not in our research area (transport), and probably also not in most other areas. A search (8 June 2022) in SCOPUS with the search string ('automated searching' AND 'literature') revealed only three hits, and none of these was relevant. On the other hand, there are already tools available (see, for example, http://systematicreviewtools.com/index.php), and we would not be surprised if using such tools will become more popular. Although the use of AI for literature reviews is still in an early and not mature stage, the future seems promising, especially if future research will support the further development of AI for this purpose (Wagner et al., 2022). If in the future authors of LRPs use AI tools, we would recommend that the use of such tools should be included in an LRP report, and this should cover the way the tools were used, the motivation for their use and a discussion of the quality of the search and selection process, highlighting issues such as the risk of selection bias. If scientometric methods are used, for example, for clustering or for searching for citation patterns, these must be explained so that they are (more or less) reproducible.

Not all journal editors, and more specifically, not all reviewers, are convinced of the fact that the methodology of an LRP should be specified clearly so that the search and selection processes are (more or less) reproducible. If so, we would suggest that a brief paragraph is presented in the main text, and a longer version is available in an appendix or online (see, for example, Boon & van Wee, 2018). Or if you disagree with a reviewer who suggests that you do not report on the methodology used, we would advise you to contact the journal's editor to discuss the topic.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## ORCID

Bert van Wee http://orcid.org/0000-0002-0370-3575
David Banister http://orcid.org/0000-0002-9002-588X

## References

Boon, W., & van Wee, B. (2018). Influence of 3D printing on transport: A theory and experts judgment based conceptual model. *Transport Reviews*, *38*(5), 556–575. https://doi.org/10.1080/01441647.2017.1370036

Chai, K.E.K., Lines, R.L.J., Gucciardi, D.F., & Ng, L. (2021). Research screener: A machine learning tool to semi-automate abstract screening for systematic reviews. *Systematic Reviews*, *10*(1), 93. https://doi.org/10.1186/s13643-021-01635-3

Le, H.T.K., Carrel, A.L., & Shah, H. (2022). Impacts of online shopping on travel demand: A systematic review. *Transport Reviews*, *42*(3), 273–295. https://doi.org/10.1080/01441647.2021.1961917

Martín-Martín, A., Thelwall, M., Orduna-Malea, E., & Delgado López-Cózar, E. (2021). Google Scholar, Microsoft Academic, Scopus, Dimensions, Web oF Science, and OpenCitations' COCI: A multidisciplinary comparison of coverage via citations. *Scientometrics*, *126*(1), 871–906. https://doi.org/10.1007/s11192-020-03690-4

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D.G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ*, *339*(jul21 1), b2535. https://doi.org/10.1136/bmj.b2535

Singh, V.K., Singh, P., Karmakar, M., Leta, J., & Mayr, P. (2021). The journal coverage of Web of Science, Scopus and Dimensions: A comparative analysis. *Scientometrics*, *126*(6), 113–5142. https://doi.org/10.1007/s11192-021-03948-5

Van Wee, B., & Banister, D. (2016). How to write a literature review paper? *Transport Reviews*, *36*(2), 278–288. https://doi.org/10.1080/01441647.2015.1065456

Wagner, G., Lukyanenko, R., & Paré, G. (2022). Artificial intelligence and the conduct of literature reviews. *Journal of Information Technology*, *37*(2), 209–226. https://doi.org/10.1177/02683962211048201