University of Groningen

Literature practices

Klitzing, Nikolai; Hoekstra, Rink; Strijbos, Jan-Willem

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in University of Groningen/UMCG research database](#)

emerald**insight**

Journal of Documentation

Literature practices: processes leading up to a citation
Nikolai Klitzing, Rink Hoekstra, Jan-Willem Strijbos,

Article information:

Users who downloaded this article also downloaded:

university of groningen

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

Literature practices: processes leading up to a citation

Nikolai Klitzing, Rink Hoekstra and Jan-Willem Strijbos
University of Groningen, Groningen, The Netherlands

Abstract

Purpose – Literature practices represent the process leading up to the citation of a source, and consist of the selection, reading and citing of sources. The purpose of this paper is to explore possible factors that might influence researchers during this process and discover possible consequences of researchers' citation behaviours.

Design/methodology/approach – In this exploratory study, various factors which could influence literature practices were explored via a questionnaire amongst 112 researchers. Participants were first authors of articles published in 2016 in one of five different journals within the disciplines of experimental psychology, educational sciences and social psychology. Academic positions of the participants ranged from PhD student to full professor.

Findings – Frequencies and percentages showed that researchers seemed to be influenced in their literature practices by various factors, such as editors suggesting articles and motivation to cite. Additionally, a high percentage of researchers reported taking shortcuts when citing articles (e.g. using secondary citations and reading selectively). Logistic regression did not reveal a clear relationship between academic work experience and research practices.

Practical implications – Seeing that researchers seem to be influenced by a variety of factors in their literature practices, the scientific community might benefit from better citation practices and guidelines in order to provide more structure to the process of literature practices.

Originality/value – This paper provides first insights into researchers' literature practices. Possible reasons for problems with citation accuracy and replicating research findings are highlighted. Opportunities for further research on the topic of citation behaviours are presented.

Keywords Social sciences, Citation, Norms, Academic shortcuts, Literature practices, Secondary citation

Paper type Research paper

1. Introduction

The social sciences are currently criticised for their limited success in reproducing the results of studies, which ultimately has led to the coining of the term "replication crisis". This topic has sparked a variety of discussions, blog posts and journal publications starting with Ioannidis (2005) reminding the research community about the importance of reproducible findings. Later, the Open Science Framework conducted a large-scale replication attempt of multiple major psychology studies (Open Science Collaboration, 2015), the result of which indicated that a large quantity of the original results could not be replicated.

The replication crisis presents a threat and an opportunity at the same time. While it can possibly lower the trust in and within the scientific community, it also offers a point of reflection and chance for improvement of the scientific mechanisms at hand. This ultimately might help form a more trustworthy and better functioning community. Since science is a communal effort, trust in each other's abilities is essential to its proper functioning. To assure good quality, there are mechanisms in place to maintain the high standards that we have set for ourselves, such as peer review and specifying in our papers how we obtained our information. How we cite and acknowledge each other specifically, however, has to our knowledge not been investigated in the social sciences so far, though there do exist theories

The authors declare no conflict of interest.

regarding why we cite each other (see Tahamtan and Bornmann, 2018; Bornmann and Daniel, 2008 for a more extensive look on this topic). Tahamtan and Bornmann named three major citation theories: normative citation theory (Merton, 1973; as cited in Tahamtan and Bornmann (2018)), social-constructivist theory (e.g. Gilbert, 1977) and a more recent theory by Nicolaisen (2007). Normative citation theory is centred around the assumption that researchers mainly cite to give and receive credit, while in social-constructivist theory citations are seen more as a tool to convince readers of the point the author is trying to make according to Gilbert (1977). Nicolaisen's theory relies on the handicap principle (Zahavi and Zahavi, 1999) and roughly states that researchers cite honestly to a certain degree as not doing so would pose a threat to their own credibility if it would be detected.

Generally speaking, the main way that we as researchers acknowledge each other's work and the community at large is through citing each other's work if we believe the normative account of citation. The path leading up to a citation, however, includes various steps such as selection of sources to cite, reading them and finally citing them. All of the steps are also prone to human error. For example, researchers have problems with interpreting commonly used statistical findings (e.g. Hoekstra *et al.*, 2014), and make mistakes when citing a source (e.g. Evans *et al.*, 1990). It is therefore also possible that they misinterpret results of another source. So far, citations and the processes leading up to them, as well as the influence of the citation process on theory building and the replicability of studies, seem to have been of little concern to the community, if the amount of literature on the topic is any indication. This problem has been acknowledged in the medical sciences, where issues with accuracy of citations are quite prevalent (e.g. Lukic *et al.*, 2004) and might consequently lead to serious distortions (e.g. Engber, 2017; Rekdal, 2014). Yet, accuracy of citations and the possible negative consequences of inaccuracies are relevant to not only the medical but to all sciences. However, the social sciences might be especially sensitive as they rely heavily on theoretical frameworks and have had difficulty reproducing some results when testing those theories (e.g. Open Science Collaboration, 2015).

The focus of this exploratory study will be on identifying factors that can influence researchers' practices when creating citations in the social sciences. Furthermore, whether and how researchers use shortcuts (e.g. secondary citations) while dealing with the works of colleagues will be investigated. Finally, it will be examined whether academic work experience influences researchers' practices. These issues will be investigated by exploring the steps of selection, reading and citing of sources, all of which can be summed up under the term literature practices.

1.1 Literature practices
The process leading up to a citation requires considerable time and effort. First, one has to search for sources and decide whether they might be relevant to one's work. Second, the source of possible relevance has to be read in order to determine whether it could be used and in order to fully understand its content. Third, after reading its content the source can, for a variety of reasons, be cited in the context of one's own current manuscript. This entire process is referred to as "literature practices" throughout the remainder of this paper and describes the overall process of dealing with colleagues' work in the context of writing a scientific manuscript.

1.2 Factors that influence researchers' literature practices
1.2.1 Norms and motivation to cite. In general, guidelines on citing are quite scarce. The APA manual provides advice on the formatting and some general statements about when one should cite and with what purpose (American Psychological Association, 2010, p. 169). However, the main responsibility for deciding when and why a citation is needed is up to the individual researcher and the research community as a collective. Nevertheless, some studies on researchers' citation behaviours have appeared. Most notably the review studies by

Bornmann and Daniel (2008), and more recently Tahamtan and Bornmann (2018) which provide a good overview on the topic of citing behaviours, including researchers' motivation to cite, and the debate around normative citing behaviours of researchers. They also introduced three major citation theories. According to Bornmann and Daniel, the evidence is mixed as to the reasons researchers cite a source. They further presented various motivations for citing (e.g. confirming previous work, providing background information and paying homage) which are not always in line with the norms (e.g. as stated in the APA manual).

Furthermore, Garzone and Mercer (2000) developed a scheme to distinguish citations, which was adapted by Bornmann and Daniel (2008) into a typology with eight different types of citations and their prevalence estimates. This typology is based on identifying the functions of each citation within the text (e.g. perfunctory, methodological and affirmation). While Bornmann and Daniel concluded that citing is a multidimensional and highly complex process that is influenced by scientific and non-scientific factors alike, their typology does not consider some of the non-scientific factors such as the influence of journal editors on researchers' citing behaviour.

1.2.2 Familiarity with articles. Researchers might rely heavily on sources they are already familiar with before the start of the literature search. This is not a bad thing *per se*, but it might become problematic if a researcher over-relies on familiar sources and their reference lists, and does not search for and include new, more recent sources. Ultimately, this might result in selection bias where newer sources will be disadvantaged, which is in line with the trend of older sources being cited again and again (Verstak *et al.*, 2014). Even though we nowadays have modern technology to help us search and select literature, we also have a higher responsibility of actually searching, reading, selecting and ultimately being familiar with a continuously increasing body of literature as data by Boyack *et al.* (2018) show.

Another problem with reusing known sources is that researchers will be less prone to re-read a source. Instead, they might rely on a previously published manuscript of theirs where they already paraphrased the source. It is also possible that they will cite the content of the source from memory. In both cases, the content of the original source might be distorted through reinterpretation or recall. Although familiarity with a source might provide a good baseline for writing a manuscript and starting the search for more sources, it might be hazardous when researchers rely too much on prior used sources and their interpretations.

1.2.3 Publication process. One of the reasons for citing a journal article can be that the researcher has a certain journal in mind to publish in, and therefore includes other publications from that same journal. Furthermore, editors might suggest sources that they would like to be included in the researcher's manuscript (Falagas and Alexiou, 2008, p. 224). Although these articles will probably be related to the content of the manuscript, the problem is the incentive behind including the source – akin to citing articles from the target journal. It is not so much about the content of the source anymore but about the inclusion of one specific source because of the journal in which it has been published – either self-selected or upon instigation by the editor. As a consequence, it likely leads to redundant citations. Bornmann and Daniel (2008) estimated the prevalence of redundant citations, based on a review of literature on the topic, to be between 10 and 50 per cent.

1.2.4 Academic work experience. The number of years one has spent actively working as a researcher can have an influence on researchers' practice. A reason for this might be a change in workload due to a higher academic position. Furthermore, the amount of training one has received on literature practices also contributed to the academic experience one has and therefore influence researchers' behaviours.

1.2.4.1 Workload and academic position. Perceived workload probably differs between researchers with varying academic work experience; typically reflected by their academic position. A high workload might, for example, influence how much time researchers allocate to reading entire sources or parts thereof. Therefore, years of academic work experience

might have an impact on literature practices, assuming that workload increases with the academic position one is holding. Choices about allocation of time due to a high workload might be a reason for why researchers take shortcuts, such as citing secondary sources and only reading parts of sources they cite. One would thus have to determine whether the current – workplace affords researchers sufficient time to properly take care of all of their tasks – including proper and accurate literature practices.

1.2.4.2 Training. As there are few guidelines on citing behaviour, this raises the question of whether researchers are properly trained on when and why to cite. This might also apply to the steps of selection and reading leading up to a citation. It is unlikely that researchers are fully aware of all the sub processes involved in proper literature practices if there is no proper training in citation norms and lack of clear guidelines for when and why to cite. Furthermore, a lack of awareness of the impact that taking shortcuts might have on the theoretical framework or theory building, in general, might partly explain why researchers are taking shortcuts in the first place.

1.2.5 Accessibility. A final factor that might influence which articles researchers cite and how they cite it, is whether the article is accessible to them. Researchers are still partly dependent on the subscription database of their own institution as it determines whether they have access to an article or not. With more and more articles being published open access and institutions being subscribed to a large number of journals, not having access hopefully will become a problem of the past eventually. Currently, however, it might nonetheless have an influence on whether researchers ultimately read and cite the original source or not.

1.3 Missteps in literature practices: shortcuts

All the prior mentioned factors might contribute to researchers engaging in shortcuts. The term "shortcuts" in the context of this manuscript refers to behaviours researchers might engage in, which deviate from the expected norms, especially when it comes to reading sources. These shortcuts include selective reading of sources one cites, and/or relying on secondary sources for information about a primary source (possibly without properly citing it). Researchers might use such shortcuts more often than they should (e.g. Goldberg *et al.*, 1993; Lukic *et al.*, 2004; Simkin and Roychowdhury, 2002).

1.3.1 Secondary citations. One commonly used shortcut is an over-reliance on secondary citations (e.g. Lukic *et al.*, 2004; Goldberg *et al.*, 1993). Rekdal (2014) described four scenarios for secondary citations: neither source is cited; it is only indicated that the information is originally from another source but obtained via a secondary source; only the secondary source is cited without citing the primary source; and the primary source is cited but the information is actually obtained via a secondary source. All scenarios have in common that the original source has not been read, and therefore the content has not been checked. However, according to APA guidelines, citing a source presupposes having read the source (American Psychological Association, 2010, p. 169). Furthermore, not reading the entire original source and relying on secondary sources instead might lead to what Rekdal (2014) called "Academic Urban Legends". These are "research findings" that supposedly are backed up by scientific research, but they are actually "myths" created by a repeated misinterpretation and repetition of that misinterpretation without consulting the original source.

Another problem with secondary citations is that the content of a citation might be distorted by repeated re-citing. Since paraphrasing is the main tool for citing in the social sciences, repeated re-wording of the text could dissociate the content ever further away from the content of the original source. A good example within the educational sciences supporting this point, is provided by Joughin (2010) who analysed three classic studies on the influence of summative assessment on learning and noted regarding the "Up to the mark study" from 1974 the following: "While the largest number of the sample were cue-deaf, this study is often cited to support the claim that students tend to be cue-seekers or

cue-conscious" (p. 340). These kinds of inaccuracies are often referred to as citation and quotation errors, which represent mistakes in the citation itself or the content of the citation, respectively (e.g. Eichorn and Yankauer, 1987). Citation errors occur with a relatively high frequency (around 20 per cent) in the medical sciences (e.g. Evans *et al.*, 1990; Fenton *et al.*, 2000; Lukic *et al.*, 2004) and might partly be caused by relying on secondary citations. It seems thus that the use of secondary citations (whether indicated or not) might also be quite prevalent in the social sciences and possibly have similar negative consequences.

Finally, the use of secondary citations might also cause problems for the use of citation metrics as a quality measure. If citations are not properly attributed, then some researchers might attain more credit than they deserve, whilst others do not receive any credit even though they should. While citation metrics as a quality measure have been criticised (e.g. Hicks *et al.*, 2015), they are still widely used to assess the quality of journals and to some extent the quality of researchers. A high reliance on secondary citations could furthermore impact the accuracy of the metrics.

1.3.2 Selective reading. Researchers quite frequently refrain from reading the original source (Simkin and Roychowdhury, 2002) or only selectively read parts of the source. Selective reading might not be as problematic as citing a secondary source as the person who cites the source at least has read the original section they are referring to, making the citation less prone to error. Nonetheless, the part that is being cited can be out of context this way as the overall message of the source that is being cited might differ from how it is interpreted by the person citing it. Furthermore, selective reading can be related to citing a secondary source as a researcher might, for example, read about a certain source in a journal article and decide to cite it, but only consults the abstract to determine what the source generally is about. This could be the case for any section of the source that the researcher deems important and decides to cite. The researcher thus only reads parts of a source that is then afterwards cited. The APA guidelines clearly state that one should have read the cited source (American Psychological Association, 2010, p. 169), which could be interpreted as having read the entire cited source. However, this can be challenging given that authors likely need to read more sources than they cite in each manuscript. Although it is hard to make an estimate of how long it takes on average to read a source in its entirety (there is probably a lot of variance between people and between sources), it seems safe to assume that reading all sources completely will be rather time-consuming.

Reading only parts of a source could take the cited information out of context. This might result in missing or misinterpreting important points that the authors make. To our knowledge there is no research that has investigated the direct consequences of selectively reading an entire source when citing it; our knowledge about the impact of such practices is hence limited.

In the context of researchers reading sources selectively, two other issues arise. First, if researchers read selectively, which parts do they select, read and cite. Second, how good is their understanding of the source if they have read it partially (or not at all).

1.4 The present study
The current study aims to obtain an overview of the process leading up to a citation in the social sciences. This will be done by investigating the steps of selecting, reading and citing literature while writing a paper (i.e. literature practices) and researchers literature practices and their work environment (workload and training) in the context of one of their recently published journal articles. Furthermore, the possible impact of researchers' work experience and prior training as well as their academic position on their literature practice will be assessed. The main research questions are as follows:

RQ1. Which factors guide/influence researchers' behaviour concerning the selection, reading and citing of literature?

RQ2. How do researchers make use of shortcuts when dealing with the work of others?

RQ3. Does academic work experience (years of research experience, training in literature practices) influence literature practices?

We hypothesise that factors such as workload and the social environment at large (e.g., training, publication process) influence researchers' literature practices. Furthermore, it is hypothesised that researchers regularly use shortcuts, such as selectively (or not) reading entire sources they cite and relying on secondary citations, while dealing with scientific literature. Additionally, it is expected that literature practices will vary with work experience.

2. Method

A multi-study approach was adopted. First, semi-structured interviews were conducted with a small sample of researchers from the local university to pilot questionnaire items prior to the main study. Second, as part of the main study the slightly modified questionnaire items were sent out to a sample of researchers who had recently published a journal article.

2.1 Pilot study

2.1.1 Participants. Ten researchers from the faculty of behavioural and social sciences at the authors' institution participated. They were selected on the basis of their specialisation, academic position and having a recent publication as a main author. The main author criterion was chosen because it was assumed that the main author is the primary person involved in reading sources and ultimately writing the manuscript. Initially 14 researchers were contacted of which 10 agreed to participate: 4 from the field of social psychology, 3 from clinical psychology and 3 from educational sciences. Three were PhD students, two assistant professors, three associate professors and two full professors.

2.1.2 Material. For the semi-structured interviews, questions were developed by the main author in cooperation with the second and third authors. The final questionnaire used in the interview consisted of 35 main questions some of which included one or multiple follow-up questions.

2.1.3 Procedure. The researchers were contacted via e-mail and asked whether they would be willing to participate in an interview regarding the process leading up to a recent journal article they first-authored. The interviews took place in the researcher's own office or a different room at the faculty. Informed consent was obtained prior to the interview. Each interview lasted around 30 min and was recorded. Although the interview was semi-structured with a set of pre-formulated questions, the main author asked follow-up questions if necessary. After each interview, the researcher could ask questions and provide feedback on the questions.

2.1.4 Results. All researchers seemed comfortable with answering the questions. However, six questions were dropped because they were considered difficult to answer or only regarded suitable for an interview setting. For some questions, changes in wording were made to improve comprehensibility and/or to tailor individual questions to the format of the main study. In the end, two questions were added. One to obtain specific information on the type of journal article that researchers were referring to, and one to obtain a percentage of familiarity with sources used.

2.2 Main study

2.2.1 Participants. Participants were researchers from three disciplines (educational sciences, social psychology and experimental psychology) who were the main authors of a journal article published in 2016 in one of five journals (*Personality and Social Psychology*

Bulletin, *Journal of Personality and Social Psychology*, *Journal of Experimental Psychology: General*, *Journal of Educational Psychology* and *American Educational Research Journal*). This resulted in 438 main authors. The current e-mail address was not identifiable for 11 authors, and another 13 e-mails bounced for unknown reasons, leaving 414 possible participants. In total, 120 researchers responded. Three did not provide consent and were excluded. Another five participants gave consent but did not answer any question and were therefore also excluded. Finally, 11 participants dropped out at some point while answering the questions and did not finish all of them. Nevertheless, it was decided to use all available information and not exclude any participants on the basis of non-response to one or multiple items. This resulted in 112 participants who answered at least a part of the questions, translating to a response rate of roughly 27 per cent. The respondents had an average academic work experience of 11.7 (SD = 7.6) years and held positions ranging from PhD student to full professor (see Table I).

2.2.2 Materials. A questionnaire was developed re-using a slight modified version of the questions used in the pilot. Modifications were based on the feedback by the interviewees and the main authors' analysis of the interviews. Qualtrics was used to distribute the questionnaire and to enabe respondents to answer more intuitively, by for example adding sliders for indicating percentages. The questionnaire consisted of 21 questions about researchers' literature practices during the time they spent working on the recently published journal article, and 21 questions about their general behaviours and experiences concerning literature practices.

2.2.3 Procedure. The first and second authors created a database of all main authors who published a journal article in 2016 in the five selected journals, including their last name, article title, journal and e-mail address. The database was used to send personalised e-mails via Qualtrics, inviting the authors to complete the questionnaire and answering some questions about their recently published journal article.

2.2.4 Analyses. Most data were analysed via descriptive statistics and percentages to provide an overview of general tendencies. Furthermore, the responses to open questions were analysed using thematic analysis to identify response patterns. Finally, logistic regression was used to analyse possible relationships between years of work experience and researchers' answers to questions indicating whether they had cited secondary sources and had read all the sources they cited in their entirety.

3. Results

Since not all researchers answered all questions or finished the questionnaire, sample sizes vary per question. For the purpose of completeness of information, the total number of valid respondents (*n*) is reported for each individual analysis. Furthermore, some of the questions had multiple answer possibilities and therefore the percentages do not add up to 100 per cent. Whenever this applies it is mentioned for the respective analysis. Additionally,

Position	Frequency (*n*)	%
PhD	15	14.9
Postdoc	29	28.7
Assistant professor	27	26.7
Associate professor	14	13.9
Full professor	11	10.9
Other	5	5.0
Note: *n* = 101		

Table I.
Frequency of academic positions in the sample

three questions were coded after the data collection using thematic analysis. For the current study, only those items relevant to our research questions were used (12 out of 21 questions specifically about the recent journal article; 14 out of 21 questions about general behaviour and experiences).

3.1 Factors that influence researchers' behaviours

3.1.1 Norms and motivation to cite. Researchers reported to include an average of 28.6 per cent (SD = 25.87, $n = 102$) sources to show expertise on the topic they were writing about, and an average of 76.4 per cent (SD = 22.75, $n = 104$) sources to support an argument they were trying to make. In general, researchers mainly described the function of citations as being to support arguments (44.9 per cent) or provide background information on the topic and embedding their own work into their research framework (62.9 per cent; see Table II). When asked about their personal reasons for using citations – in contrast to the general function of citation – providing background information on the topic was mentioned less frequently (46.4 vs 64 per cent), whereas showing expertise (19 vs 10.1 per cent) and pragmatic reasons (16.7 vs 1.1 per cent) were mentioned more often (see Table II).

3.1.2 Publication process. Most researchers relied on Google Scholar for finding relevant articles (84.3 per cent), followed by EBSCOhost (28.7 per cent) and reference lists of familiar sources (28.7 per cent) (see Table III). When asked about the influence of the journal to which they planned to submit, some researchers (22.5 per cent; 23 out of 102) stated that they had included sources previously published in that journal. Additionally, editors of a journal had asked 53.9 per cent (55 out of 102) of researchers to include a source, while 2.9 per cent (3 out of 102) were asked to exclude manuscript source. About half of the researchers (45.1 per cent; 46 out of 102) reported to have not been asked by the editor to either include or exclude sources. When asked to include or exclude sources by editors, all researchers

Answer	General function		Personal reasons	
	Frequency (n)	%	Frequency (n)	%
Support arguments	42	47.2	35	41.7
Give credit	16	18.0	18	21.4
Background/Embed	57	64.0	39	46.4
Theorising	5	5.6	3	3.6
Show expertise	9	10.1	16	19.0
Pragmatic	1	1.1	14	16.7
Other	6	6.7	6	7.1
Note: Multiple answers possible, $n = 89$ and 84				

Table II. General function and personal reasons for citing

Answer	Frequency (n)	%
EBSCOhost	31	28.7
Google scholar	91	84.3
Specific journals	21	19.4
Snowball	31	28.7
Web of science	21	19.4
Reference program	5	4.6
Other	33	30.6
Note: Multiple answers possible, $n = 108$		

Table III. Main sources for finding relevant articles

reported to have done so; i.e. add or remove at least one of the suggested sources if asked to include or exclude multiple sources.

3.1.3 Familiarity. In the sample, 46.4 per cent of the researchers indicated being already familiar with the sources they included before searching for new sources. Of the 112 respondents, 49.1 per cent indicated that they were not familiar with any of the sources they included before starting to search for any sources (see Table IV). When asked to indicate the percentage of sources they were familiar with before starting their search, they responded with an average of 61.5 per cent (SD = 21.55 per cent, $n = 111$). Most researchers reported they had already done earlier work into the same direction as their recently published article (70.5 per cent; 79 out of 112).
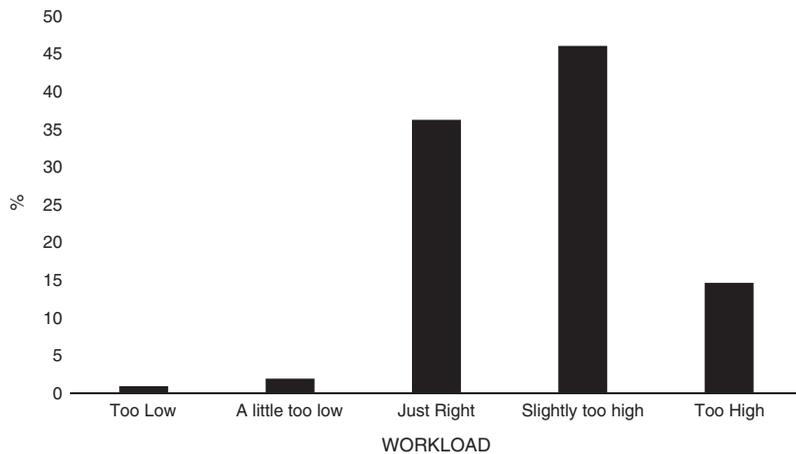
3.1.4 Academic work experience. 3.1.4.1 Workload and academic position. The perceived workload seems to increase throughout a research career, with 77.5 per cent (79 out of 102) of researchers reporting a higher workload now than earlier in their career. Additionally, 46.1 per cent (47 out of 102) of researchers perceived their workload to be somewhere between just right and too high (slightly too high) (see Figure 1).

Regarding the effort needed to understand individual sections, the Method and Results sections were mentioned most often (54.2 and 58.3 per cent, respectively). When asked which individual sections required the most time reading, the Results (54.2 per cent), Method (50 per cent) and Introduction (49 per cent) sections were mentioned most often (see Table V).

Table IV.
Were you already familiar with articles included in the reference list before you started the search procedure for the recent paper?

Answer	Frequency (n)	%
Yes	52	46.4
No	55	49.1
Do not know	4	3.6
Partially	1	0.9

Note: $n = 112$

Figure 1.
Perceived workload of researchers

Note: $n=102$

Logistic regression analysis was used to determine whether years of academic work experience influenced literature practices such as citing secondary sources or reading all included sources. Interaction terms were added to the equation to test the assumption of linearity. In both cases, there was no clear violation of linearity. The individual variables were not significant: citing secondary source ($n = 101$) resulted in $p = 0.643$; and reading all the sources ($n = 101$) in $p = 0.674$. This indicates that there was no clear evidence of researcher's academic work experience influencing the use of shortcuts in our sample.

3.1.4.2 Training. Of the 100 researchers who answered the question about citation training, 43 reported that they had received formal training on when and why to cite at some point in their career, whereas the other 57 reported that they had not received any formal training.

3.1.5 Access and importance. About two-thirds of the researchers (64.4 per cent; 67 out of 104) reported that they searched digital documents for keywords. Getting access to sources was most of the time not a problem for them (70.1 per cent), but if they encountered difficulties most researchers relied on their local library services (13.4 per cent) (see Table VI).

Researchers were also asked to indicate which sections of a source they find the most and which the least important. The Method and Results sections were considered the most important sections (61.6 and 69.7 per cent, respectively). Additionally, 54.9 per cent of researchers considered the Discussion to be the least important section (Table VII).

Answer	Effort frequency (*n*)	%	Time frequency (*n*)	%
Abstract	5	5.2	13	13.5
Introduction	24	25.0	47	49.0
Method	52	54.2	48	50.0
Results	56	58.3	52	54.2
Discussion	11	11.5	35	36.5
Note: Multiple answers possible, $n = 96$				

Table V.
Sections that on average take the most effort and time to read

Answer	Frequency (*n*)	%
No	68	70.1
Original author	7	7.2
Colleagues	6	6.2
Library	13	13.4
ResearchGate	2	2.1
Other	8	8.2
Note: Multiple answers possible, $n = 97$		

Table VI.
Did you run into problems with accessing articles? And if so, how did you solve them?

| Section | Most important | | Least important | |
	Frequency (*n*)	%	Frequency (*n*)	%
Abstract	31	31.3	20	22.0
Introduction	40	40.4	22	24.2
Method	61	61.6	10	11.0
Results	69	69.7	9	9.9
Discussion	31	31.3	50	54.9
Note: Multiple answers possible, $n = 99$ and 91, respectively				

Table VII.
Most and least important sections on average

3.2 Missteps in literature practices: shortcuts

Citing secondary sources without reading or with only partially reading them was practiced by around one-third (32 per cent; 33 out of 103) of researchers in the sample, whereas 68 per cent (70 out of 103) reported not having cited any secondary sources in their journal article. Furthermore, when asked about the extent that they had read all the sources cited entirely, 23.1 per cent (24 out of 104) claimed that they had and 76.9 per cent (80 out of 104) that they had not read done so. Additionally, researchers reported that they focussed mainly on the Results section when reading sources (83.5 per cent), followed by the Method section (68 per cent) and abstract (67 per cent) (see Table VIII).

The majority of researchers (85.1 per cent; 86 out of 101) reported that they never received a complaint from a colleague feeling inaccurately cited by them. Nonetheless, 49.5 per cent (50 out of 101) reported that they felt inaccurately cited by colleagues at some point in their career. However, only few pointed this out to the colleague whom they felt had cited them inaccurately (18.5 per cent; 10 out of 54).

4. Discussion

This exploratory study investigated possible factors that influence researchers' literature practices in the process of publishing a journal article. It was hypothesised that researchers sometimes rely on shortcuts in the process of citing sources. Furthermore, it was expected that researchers' years of work experience would influence their literature practices and consequently the extent to which they use shortcuts.

4.1 Factors that influence researchers' literature practices

4.1.1 Norms and motivation to cite. Inquiry into researchers' personal reasons to cite in comparison to the general function of citations revealed that personal reasons play a role in citing certain sources. For example, when asked about their personal reasons for citing, researchers mentioned more often that they cited to show expertise, compared to when they were asked about the general function of citing. Furthermore, researchers also stated pragmatic reasons (e.g. having to add a citation because the system demands it or to please reviewers) when asked about their personal motivation for citing more often compared to when they were asked about the general function of citations. Pragmatic reasons are not part of Bornmann and Daniel's (2008) typology. These can only be inferred from direct inquiry rather than from analysing citations. Since pragmatic reasons seem to play a role in personal motivations for citing, these reasons for citing should be further investigated. Moreover, future research on literature practices should be expanded with direct measures and not solely rely on indirect measures, as this allows for better insight into researchers' personal motivation to cite.

4.1.2 Publication process. Researchers indicated that journal editors regularly suggest sources that they would like to be included. In almost all cases, the researchers followed the suggestions. This behaviour by editors is considered a way to improve the impact factor of

Answer	Frequency (n)	%
Abstract	69	67.0
Introduction	49	47.6
Method	70	68.0
Results	86	83.5
Discussion	39	37.9

Table VIII.
Main focus of reading   **Note:** Multiple answers possible, $n = 103$

the journal, if the suggested source was indeed published in the same journal (Falagas and Alexiou, 2008). It has to be noted that it was not checked whether suggested articles were indeed articles from the same journal as the one that the manuscript was submitted to. The reason for suggesting the article could thus also be that the article was thought to be useful to the authors of the manuscript.

Another potential problem with suggesting articles, from the same journal or not, is that the authors of the submitted manuscript might not take the risk and ignore the editor's suggestion, presumably because they are afraid that their manuscript will rejected otherwise. This could have the unwanted consequence that researchers do not read the source suggested by the editor thoroughly. Instead, they might merely cite it without much knowledge about the exact content and suitability of the source. Further inquiry into how researchers deal with sources suggested by editors could result in a better understanding of the impact of publication process on literature practices.

4.1.3 Familiarity. How familiar researchers are with sources influences which sources are selected. Many researchers had done previous work into the same direction as the recent journal article that they were asked about and were therefore already familiar with one or multiple sources. This, in turn, might influence reading and selection behaviours, as the need to search and select sources is lower, and researchers might not re-read entire sources they already cited. Additionally, errors might occur in citing of familiar sources if the recall of information is erroneous. Closer examination of researchers' re-reading behaviours of familiar sources is needed to identify whether this can have negative consequences.

4.1.4 Academic work experience. 4.1.4.1 Workload and academic position. Researchers indicated that their workload increases throughout their career and about half of them perceived their workload as slightly too high. Furthermore, researchers varied in their number of years of work experience and held different academic positions. Nonetheless, the difference in years of work experience did not seem to have a significant influence on the use of shortcuts within our sample. This was unexpected given the fact that perceived workload seems to steadily increase – according to the researchers themselves – as one moves up in academic position. A non-significant result can mean one of two things: Either there is no population effect, or there is one but the study was not powerful enough to detect it. In the context of this study, this could mean that perhaps the differences in years of work experience are not large enough to have resulted in a significant impact on behaviours. Alternatively, it could be that we were not able to detect a significant relationship between the two variables due to a lack of power of this study. However, it is also possible that the actual difference in workload is not so different across academic positions as we worked with self-reported, perceived workload which might be subjective to bias. Another explanation regarding the non-significant outcome of the analysis could be that researchers also become more experienced when it comes to time management, thereby becoming more efficient in their work and the delegation of tasks. Since a majority of the researchers indicated that they did engage in behaviours such as selectively reading; however, it is possible that most researchers engage in these behaviours independent from their work experience. Future studies could directly compare different work environments and their specific impact on literature practices to investigate the role of academic work experience more specifically.

4.1.4.2 Training. Less than 50 per cent of the researchers indicated that they ever received formal training on when and why to cite. This is also partly reflected by college and university professors in the USA having different conceptions about what paraphrasing actually should look like (Roig, 2001). Consequently, researchers might not be (sufficiently) aware of the best practice(s) when it comes to literature practices and possibly need further training. The lack of training could also be one of the reasons why researchers in the present sample engaged in shortcuts such as citing secondary sources.

4.1.5 Access and importance. Most researchers considered the Method and Results as the most important sections, whereas notably, the Discussion was deemed to be the least important. This corresponds with researchers indicating that they mainly focussed on the Method and Results sections. Accessibility of sources also seemed to be a factor as around 30 per cent of researchers had difficulties accessing articles initially, though a lot of them seemed to able to ultimately get access by contacting their local library services, colleagues or the original author. Therefore, problems with initially accessing an article do not serve as a sufficient explanation for why researchers might rely on secondary citations, although they might still have some influence on such practices. Further studies investigating the link between these two variables are needed, also in connection with possible implications for the usefulness of publishing open access.

4.2 Missteps in literature practices: shortcuts

Most researchers appeared to use two shortcuts when asked directly about them: inclusion of secondary citations and selective reading of the source they cite. A reason for using such shortcuts might be that researchers merely act out of convenience – which has led to the term "lazy author syndrome" (e.g. Gavras, 2002) – when they opt to cite secondary sources.

4.2.1 Secondary citations. The use of shortcuts such as relying on secondary sources for citations can lead to serious errors, as they might result in inaccuracies of the citation. This roughly corresponds to around half of the researchers in the sample reporting that they felt cited wrongly at some point in their career. While that link has not been directly investigated in the present study, quotation accuracy and rate of secondary citations are often investigated together (e.g. Goldberg *et al.*, 1993; Lukic *et al.*, 2004). This makes sense as citing secondary sources can distort the message of the original source and might eventually lead to an erroneous interpretation of the original source. Additionally, the APA guidelines state that one should only fall back on using secondary citations when the original source is not available (American Psychological Association, 2010, p. 178). Since access to sources was not problematic for most researchers in the present study, they could have complied with the APA guidelines. However, the results of the current study show the opposite with around one-third of the researchers indicating to rely on secondary citations. One reason for this might be the previously discussed problems with accessing articles initially. However, the exact reasons for why so many researchers rely on secondary citations remain unclear and signal the need for further investigation.

4.2.2 Selective reading. While exploring why researchers did not read all cited sources entirely, it became apparent that most researchers focussed on the Results and Method sections while reading. Since these sections were also regarded as the most important ones, it might be the case that the researchers only focus on the sections they deem the most relevant, and then cite the source without having read it completely. Additionally, time constraints due to a heavy workload might play a role as well, because researchers might not deem it possible to read all the sources entirely. The role of time constraints and the impact of selective reading need to be further explored in future studies.

4.3 Problems with terminology

While investigating literature practices, it became apparent that the terminology used is rather ambiguous. Usually authors only describe what is meant by the terms quotation accuracy and quotation error (e.g. Evans *et al.*, 1990; Eichorn and Yankauer, 1987). However, they do not specify whether they apply these terms only to direct quotations, paraphrasing of sources or both. Researchers in this field should be more specific in what is meant by the term quotation accuracy to avoid conceptual and methodological confusion and to make the topic more accessible to the broader scientific community. One option would be to add

the term paraphrasing accuracy when referring to issues with paraphrasing. By doing so, the term quotation accuracy can exclusively be used for direct quotations. Moreover, APA provides no guidelines for the accuracy of paraphrasing and only mentions direct quotations (American Psychological Association, 2010, p. 172). This could be expanded upon by the APA in order to provide more specific guidelines to researchers.

4.4 Methodological limitations

One of the major limitations of the current study is that it relies on self-reported questionnaire data. This might result in self-serving biases on the respondents' side, leading to conservative estimates, as with, for example, the number of researchers admitting to have relied on secondary citations. However, the fact that undesirable answers were given regularly indicates that apparently researchers were willing to answer questions truthfully.

Another limitation is that many questions referred to the process leading up to a publication. Due to the often-long time periods between writing and actually publishing a paper, the time period that was of interest was sometimes over two years ago. Some of the participants mentioned this to be a problem with regard to accurately answering the items for which they had to provide estimates.

Finally, only a fraction of the contacted authors actually responded (resulting in a 27 per cent response rate), which is reflected by the relatively small sample size ($n = 112$). It is therefore possible that authors who were not willing to answer possibly uncomfortable question because they perceived their own answers as undesirable. Therefore, the results reported here might be even more positive than in reality when it comes to literature practices. This also means that we do not know to what extent the final sample used for analysis is actually representative.

4.5 Practical implications

Given the extent to which the researchers reported the use of shortcuts, it is advisable to train current and future students more extensively in literature practices and specifically the possible harmful effects of citing secondary sources. This should consist of teaching good literature practices already at the undergraduate level to ensure that citations are at least motivated by the norms stated in the APA manual (or other discipline-specific structured guidelines). However, this presupposes specific guidelines for not only when and why to cite, but also which conditions should have been met to warrant a citation. Adding page numbers to references might furthemore help reduce the rate of secondary citations. By doing so, it will become easier to verify citations, thus the need to rely on a secondary citations will decrease (Rekdal, 2014). Finally, even though the topic still requires more investigation, editors' practice of suggesting articles to researchers could be problematic if they only suggest the articles with the intent to improve the impact factor of their own journal and therefore should be examined in more detail.

5. Conclusion

To our knowledge, this study is the first that provided an exploration of literature practices and more specifically the process leading up to a citation by asking researchers about their behaviours concerning the selection, reading and citing of sources. Literature practices deserve more attention in the research community, as improper practices likely contribute to citation, quotation and paraphrasing errors. Ultimately, literature practices also constitute a reflection of the state of research at the moment and offer one of many starting points towards finding a solution of the still ongoing replication crisis. Moreover, literature practices are not only relevant to disciplines that are currently affected by the replication crisis, but all scientific disciplines in general. Although literature practices are affected by a

multiplex of factors (such as workload, time pressure, journal editors' behaviours and individual citation practices) they require close up investigations to obtain a more accurate picture of how they specifically influence researchers' behaviours. Hopefully, that way we can ultimately ascertain the role of citations as a reliable and valid practice for acknowledging each other's work and constructing and expanding the theoretical frameworks within which we conduct and interpret our work.

References

American Psychological Association (2010), *Publication Manual of the American Psychological Association*, 6th ed., American Psychological Association, Washington, DC.

Bornmann, L. and Daniel, H. (2008), "What do citation counts measure? A review of studies on citing behavior", *Journal of Documentation*, Vol. 64 No. 1, pp. 45-80, available at: https://doi.org/10.1108/00220410810844150

Boyack, K.W., van Eck, N.J., Colavizza, G. and Waltman, L. (2018), "Characterizing in-text citations in scientific articles: a large-scale analysis", *Journal of Informetrics*, Vol. 12 No. 1, pp. 59-73, available at: https://doi.org/10.1016/j.joi.2017.11.005

Eichorn, P. and Yankauer, A. (1987), "Do authors check their references? A survey of accuracy of references in three public health journals", *American Journal of Public Health*, Vol. 77 No. 8, pp. 1011-1012, available at: https://doi.org/10.2105/AJPH.77.8.1011

Engber, D. (2017), "Bad footnotes can be deadly", *Slate*, 11 June, available at: www.slate.com/articles/health_and_science/science/2017/06/how_bad_footnotes_helped_cause_the_opioid_crisis.html (accessed 20 September 2017).

Evans, J.T., Nadjari, H.I. and Burchell, S.A. (1990), "Quotational and reference accuracy in surgical journals: a continuing peer review problem", *Journal of the American Medical Association*, Vol. 263 No. 10, pp. 1353-1354, available at: https://doi.org/10.1001/jama.1990.03440100059009

Falagas, M.E. and Alexiou, V.G. (2008), "The top-ten in journal impact factor manipulation", *Archivum Immunologiae et Therapiae Experimentalis*, Vol. 56 No. 4, pp. 223-226, available at: https://doi.org/10.1007/s00005-008-0024-5

Fenton, J.E., Brazier, H., de Souza, A., Hughes, J.P. and McShane, D.P. (2000), "The accuracy of citation and quotation in otolaryngology/head and neck surgery journals", *Clinical Otolaryngology & Allied Sciences*, Vol. 25 No.1, pp. 40–44, available at: https://doi.org/10.1046/j.1365-2273.2000.00322.x

Garzone, M. and Mercer, R.E. (2000), "Towards an automated citation classifier", in Hamilton, H.J. (Ed.), *Advances in Artificial Intelligence*, Springer, Berlin and Heidelberg, pp. 337-346, available at: https://doi.org/10.1007/3-540-45486-1_28

Gavras, H. (2002), "Inappropriate attribution: the 'lazy author syndrome'", *American Journal of Hypertension*, Vol. 15 No. 9, p. 831, available at: https://doi.org/10.1016/S0895-7061(02)02989-8

Gilbert, G.N. (1977), "Referencing as persuasion", *Social Studies of Science*, Vol. 7 No. 1, pp. 113-122.

Goldberg, R., Newton, E., Cameron, J., Jacobson, R., Chan, L., Bukata, W.R. and Rakab, A. (1993), "Reference accuracy in the emergency medicine literature", *Annals of Emergency Medicine*, Vol. 22 No. 9, pp. 1450-1454, available at: https://doi.org/10.1016/S0196-0644(05)81995-X

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S. and Rafols, I. (2015), "Bibliometrics: the Leiden Manifesto for research metrics", *Nature News*, Vol. 520 No. 7548, p. 429, available at: https://doi.org/10.1038/520429a

Hoekstra, R., Morey, R.D., Rouder, J.N. and Wagenmakers, E.-J. (2014), "Robust misinterpretation of confidence intervals", *Psychonomic Bulletin & Review*, Vol. 21 No. 5, pp. 1157-1164, available at: https://doi.org/10.3758/s13423-013-0572-3

Ioannidis, J.P.A. (2005), "Why most published research findings are false", *PLOS Medicine*, Vol. 2 No. 8, pp. 0696-0701, available at: https://doi.org/10.1371/journal.pmed.0020124

Joughin, G. (2010), "The hidden curriculum revisited: a critical review of research into the influence of summative assessment on learning", *Assessment & Evaluation in Higher Education*, Vol. 35 No. 3, pp. 335-345, available at: https://doi.org/10.1080/02602930903221493

Lukic, I.K., Lukic, A., Gluncic, V., Katavic, V., Vucenik, V. and Marusic, A. (2004), "Citation and quotation accuracy in three anatomy journals", *Clinical Anatomy*, Vol. 17 No. 7, pp. 534-539, available at: https://doi.org/10.1002/ca.10255

Merton, R.K. (1973), *The Sociology of Science: Theoretical and Empirical Investigations*, University of Chicago Press, Chicago, IL.

Nicolaisen, J. (2007), "Citation analysis", *Annual Review of Information Science and Technology*, Vol. 41 No. 1, pp. 609-641, available at: https://doi.org/10.1002/aris.2007.1440410120

Open Science Collaboration (2015), "Estimating the reproducibility of psychological science", *Science*, Vol. 349 No. 6251, pp. 943-951, available at: https://doi.org/10.1126/science.aac4716

Rekdal, O.B. (2014), "Academic urban legends", *Social Studies of Science*, Vol. 44 No. 4, pp. 638-654, available at: https://doi.org/10.1177/0306312714535679

Roig, M. (2001), "Plagiarism and paraphrasing criteria of college and university professors", *Ethics & Behavior*, Vol. 11 No. 3, pp. 307-323, available at: https://doi.org/10.1207/S15327019EB1103_8

Simkin, M.V. and Roychowdhury, V.P. (2002), "Read before you cite!", available at: http://arxiv.org/abs/cond-mat/0212043 (accessed 16 February 2017).

Tahamtan, I. and Bornmann, L. (2018), "Core elements in the process of citing publications: conceptual overview of the literature", *Journal of Informetrics*, Vol. 12 No. 1, pp. 203-216, available at: https://doi.org/10.1016/j.joi.2018.01.002

Verstak, A., Acharya, A., Suzuki, H., Henderson, S., Iakhiaev, M., Lin, C.C.Y. and Shetty, N. (2014), "On the shoulders of giants: the growing impact of older articles", available at: http://arxiv.org/abs/1411.0275 (accessed 14 May 2017).

Zahavi, A. and Zahavi, A. (1999), *The Handicap Principle: A Missing Piece of Darwin's Puzzle*, Oxford University Press, Oxford.

Further reading

Rekdal, O.B. (2014), "Academic citation practice: a sinking sheep?", *Portal: Libraries and the Academy*, Vol. 14 No. 4, pp. 567-585, available at: http://doi.org/10.1353/pla.2014.0025

Corresponding author
Nikolai Klitzing can be contacted at: n.f.klitzing@gmail.com