

## Survey on inadequate and omitted citations in manuscripts: a precursory study in identification of tasks for a literature review and manuscript writing assistive system

[Aravind Sesagiri Raamkumar](#), [Schubert Foo](#), and [Natalie Pang](#)

### Abstract

**Introduction.** This paper looks at the issue of inadequate and omitted citations in manuscripts by collecting the experiential opinions of researchers from the dual perspectives of manuscript reviewers and authors.

**Method.** An online survey was conducted with participation from 207 respondents who had experience of reviewing and authoring research papers.

**Analysis.** The collected data were analysed quantitatively. Descriptive and bivariate analyses were performed.

**Results.** Reviewer and author groups opined that manuscript authors fail to cite seminal and topically-similar papers, while the reviewer group indicated that authors include too few papers and cite irrelevant papers. The lack of experience was perceived as a major reason for inadequate and omitted citations, followed by lack of overall research experience and working in interdisciplinary research projects. Authors needed external assistance in finding papers for a literature review. Google Scholar was the most used system among the list of information sources.

**Conclusions.** The findings may benefit subsequent studies conducted to solve the issue of inadequate and omitted citations through process improvements and technological interventions. The findings helped in identifying three tasks for a literature review and manuscript writing assistive system. The usage preferences on information sources helped in shortlisting Google Scholar's user-interface as a basis for the user-interface design for the assistive system.

### Introduction

The scientific publication lifecycle encompasses the key

activities in academic research ([Björk and Hedlund, 2003](#)). During a literature review, researchers identify research gaps by synthesising prior studies. Subsequently, a research framework is devised, data is collected and required analyses are performed ([Levy and Ellis, 2006](#)). The lifecycle completes a single iteration with the publication of study results, followed by the next set of related studies. The sequence of activities where researchers seek information is outlined in the scientific information seeking model ([Ellis and Haugan, 1997](#)). The execution complexity of all these activities clearly differentiate experienced researchers from beginners ([Karlsson et al., 2012](#)). Research has been conducted in the past to understand the issues so as to ease the complexity of each activity in this lifecycle for novice researchers. There are two kinds of non-mutually exclusive interventions. They are *process-based* and *technology-oriented* interventions.

The importance of librarians and expert researchers has been highlighted in helping rookie researchers, from a process standpoint ([Du and Evans, 2011](#); [Spezi, 2016](#)). Some of the proposed technology-oriented interventions include search systems with faceted user interfaces for better display of search results ([Atanassova and Bertin, 2014](#)), symbiotic systems ([Gamberini et al., 2015](#)), meta-search systems ([Sturm, Schneider and Sunyaev, 2015](#)), bibliometric tools for visualising citation networks ([Chou and Yang, 2011](#)) and scientific paper recommender systems in both embedded mode ([Beel, Langer, Genzmehr and Nürnberger, 2013](#); [Naak, Hage and Aïmeur, 2008](#)) and stand-alone mode ([Huang, Wu, Mitra and Giles, 2014](#); [Küçüktunç, Saule, Kaya and Çatalyürek, 2013](#)) for catering to specific literature review search and manuscript writing tasks. There is an apparent lack of a task-based assistive system where key literature review search tasks are linked. The available options are piecemeal approaches catering to different tasks.

In the context of studies conducted to identify issues at each activity level in the lifecycle, both qualitative and quantitative methods have been employed. There has been a lack of contextual insights with a key issue taken as the central frame of reference across the activities. This type of approach is required to bind the observations across the activities. In order to identify a key issue, there is a need to look at the criticality of the consequences of the issues. The inadequacies of research studies are identified at the stage where official reviewers review the manuscripts submitted to journals and conferences. Manuscript reviewers have cited issues such as missing citations, unclear research objectives and improper research framework ([Bornmann, Weymuth and Daniel, 2009](#); [McKercher, Law, Weber and Hsu, 2007](#)). From a citation analysis perspective,

issues such as formal influences not being cited, biased citing and informal influences not being cited, have been identified as problems with cited papers in manuscripts ([MacRoberts and MacRoberts, 1989](#)). An improper literature review often leads to problems in subsequent stages of research since findings are not compared with the correct and sufficient number of prior studies. Thus, inadequate and omitted citations in the related work section of manuscripts are indicators of poor quality in research. The current study aims to investigate this issue of inadequate and omitted citations in manuscripts and the associated factors such as the different instances, reasons and the scenarios where external assistance is required.

We conducted a survey-based study for collecting experiential opinions of manuscript reviewers and authors with specific focus on the aforementioned issue. The perspectives of researchers as information seekers have been adequately explored in the context of information seeking performed during literature review. It is expected that the perspectives of researchers as manuscript reviewers and authors would provide insights from the academic writing context since it is a bottom-up approach. The findings of this study are intended to inform the design of a task-based scientific paper recommendation system which is currently under concurrent development and evaluation. The system is aimed at assisting researchers in their literature review and manuscript writing tasks. The survey covered the topics - instances of citations missed by authors, reasons for inadequate and omitted citations, tasks where authors needed external assistance and usage preferences of information sources such as academic search engines, databases, digital libraries and other academic systems. The inadequate and omitted citation instances directly relate to specific literature review recommendation tasks. For example, authors who tend to miss citing seminal papers in their manuscript would be benefitted by the corresponding remedial task where seminal papers are specifically recommended. This scenario has been addressed in few previous studies ([Bae, Hwang, Kim and Faloutsos, 2014](#); [Wang, Zhai, Hu and Chen, 2010](#)).

This is the first study to use the dual perspectives of manuscript reviewers and authors as both these roles are important in ascertaining the criticality of the different types of inadequate and omitted citations. The findings from this study will help researchers gain a holistic understanding of the issue of inadequate and omitted citations in research manuscripts. The findings will also provide empirical evidence to justify the need for developing holistic intelligent systems such as recommender systems that cater to the different search tasks performed during literature review and academic writing.

## Literature review

### Issues faced by researchers during information seeking

The consolidated Ellis information seeking model ([Ellis and Haugan, 1997](#)) has been considered as the base model for guiding studies in ascertaining the intricacies and issues in the information seeking sessions of academic researchers. Prior studies provide an overview of issues faced by researchers during information seeking. Barrett (2005) studied the information seeking habits of graduate students in a Humanities school. Findings of this study show that individual sessions of graduates tend to be idiosyncratic with projects started in a haphazard manner. Results indicate students' reliance on supervisors' guidance in finding relevant documents, thereby underlining the importance of providing intelligent systems in addition to online library services such as online public access catalogues. Barrett's study is in line with an earlier study conducted by Fidzani (1998) which highlights the graduate students' inability in using the library resources in a required manner to perform research. The assistive role required from library resources and librarians in helping researchers is a common observation by many studies ([Catalano, 2013](#); [George et al., 2006](#); [Head, 2007](#); [Wu and Chen, 2014](#)). The study conducted by Du and Evans (2011) identified characteristics, strategies and sources related to research-related search tasks of graduate students. The study indicates the difficulties faced by researchers in finding relevant information using multiple explorative search sessions in multiple sources. The nuances in handling the features provided by academic search systems are highlighted with marked differences between novices and experts. This issue is also echoed in other studies ([Brand-Gruwel, Wopereis and Vermetten, 2005](#); [Karlsson et al., 2012](#); [Yoo and Mosa, 2015](#)) where the experts' ability in carefully formulating a problem before conducting a search is highlighted as a key difference. Academic experience and searching competency are two factors that affect the confidence of researchers, thereby differentiating expert researchers from novices ([Niu and Hemminger, 2012](#)).

Studies related to exploratory search and search tasks have identified the characteristics of the search tasks undertaken by researchers and attributed the complexity of these tasks to characteristics such as uncertainty and prior knowledge of researchers ([Liu and Kim, 2013](#); [Wildemuth and Freund, 2012](#)). Directed search and passive browsing are the most-used search techniques by researchers during their information seeking sessions ([Bates, 2002](#); [Vilar and Žumer, 2011](#)), with the former

representing the search for an exact information need while the latter being more serendipitous in nature. Most of these research studies suggest remedial measures such as educating new users on information seeking best practices, involving librarians for assisting users and involving experts to assist novices so that their information seeking, searching skills and competencies improve.

## Citing behaviour in academic writing

Scientific research has a cumulative nature of building on top of prior studies for advancement. Therefore, studying citation behaviour, citation patterns and citation intents have remained important. Erikson and Erlandson (2014) have put forth a taxonomy of motivations based on the citation intent of researchers. The main categories are argumentation (support for claims), social alignment (pushing author's identity), mercantile alignment (uplifting other researchers' works) and data (evidence). Other citation motivation studies (Case and Higgins, 2000; Case and Miller, 2011; Harwood, 2009) have reported along similar lines. Even though citations are made with any of the aforementioned motivations, it should not be taken for granted that authors have read the papers that they cite. Simkin and Roychowdhury (2002) report that the majority of authors do not tend to read the papers that they cite and these are the papers with high citation counts. It is not known whether this behaviour is seen across researchers at different experience levels. A study by Oppenheim and Renn (1975) on highly cited papers identified that the repeated citations of these papers are mainly for historical reasons as an indication of continued contemporary relevance. Some of these papers may be considered seminal in the particular research area; therefore the authors may consider the citation as necessary. The impact of publication venues on citation of papers in manuscripts has been identified as a differentiating factor behind citation and non-citation of papers. MacRoberts and MacRoberts (2010) conducted a study in the context of biogeography papers. It was found that the papers from Thomson Reuters monitored journals had more citation visibility as against the other publication venues. However, this finding needs to be validated across disciplines. Interestingly, authors' intent behind citations of papers and the readers' understanding of the citation context vary considerably (Willett, 2013). Factors such as differences in primary discipline(s), interest areas and research experience can be attributed to the dissimilarity in understanding.

## Manuscript rejections

Research manuscripts' acceptance in publications is subject to scrutiny by reviewers who are experts in the particular research

areas. Peer review of manuscripts is seen as the mechanism for quality control in scientific research (Braun, 2004). Manuscripts are rejected by reviewers if the expected quality is not met. Studies have been conducted in the past looking at the reasons behind manuscript rejections. The common reasons for rejection are issues with research contributions, methodology, design, irrelevant results, and results with inconclusive evidence (Bornmann, Nast and Daniel, 2008; Bornmann *et al.*, 2009; Byrne, 2000). The rejection category 'Reference to the literature and documentation' (Bornmann *et al.*, 2009) is not regarded as a major reason for rejecting manuscripts. In different studies, this category is ranked at different positions based on the discipline and the journals studied. For instance, it is number 8 in manuscript rejection reasons in Bordage's (2001) study, number 4 in McKercher *et al.* (2007) and part of number 2 (sub-category) in Murthy and Wiggins (2002). McKercher *et al.* (2007) have outlined the list of issues in the manuscripts that have a weak literature review section. Poor referencing, lack of key citations, irrelevant and outdated citations are some of the citation-related issues identified. The studies conducted so far have used the review log of publications to arrive at conclusions.

Prior studies have contributed to the understanding of the information seeking behaviour of researchers, as well as the identification of relevant issues and other contextual factors. However, the existing research also has some limitations. While studies have collected data in the context of information seeking, the impact of the seeking/searching sessions on conducting research and manuscript writing have not been explored. There is also a lack of studies that look at the citing behaviour of researchers at different experience levels. The type of citation for which researchers have difficulty finding relevant papers is another area unexplored. Data has not been collected directly from researchers on the topic of manuscript review with emphasis on the different instances of inadequate and omitted citations, and the related reasons.

## **Aims of research**

This study intends to investigate the issue of inadequate and omitted citations in manuscripts and associated factors by addressing the following questions:

Q1: What are the critical instances of inadequate and omitted citations in research manuscripts submitted for review in journals and conferences?

Q2: Do the critical instances and reasons of inadequate and omitted citations in research manuscripts relate with the scenarios/tasks where researchers need external assistance in finding papers?

*Motivation: Q1 and Q2 will help in identifying the literature review and manuscript writing tasks that are to be considered for an assistive system.*

Q3: Identify the prominent information sources that include academic databases, search engines and digital libraries, through frequency of usage.

*Motivation: Q3 will help in identifying the system(s) which can be considered as a basis for designing the User Interface screens and selecting the display features of the proposed assistive system.*

## Methods

### Online survey instrument

In an earlier study ([MacRoberts and MacRoberts, 1988](#)) conducted for identifying author motivations for citing certain papers and omitting others, an interview-based method was used for data collection. Interviews were not considering for our study as the aims of the research were not suited to such a method. Data for this study was collected through questionnaire surveys. The questionnaires were pre-tested in a pilot study. The survey flow and questions were corrected based on the feedback of the participants for the criteria readability and comprehensibility. Since there were two types of participants in the study, the section meant for collecting data from the perspective of a reviewer was hidden for participants who identified themselves as mainly authors with no reviewing experience. Demographic details of the participants were collected at the start of the survey: data about age group, gender, highest education level, current position, parent school, primary discipline. The data pertaining to the study was collected in four main sections. Details about the different sections from the full questionnaire (provided to the reviewer group) are provided in Table 1. The questions for the segments in Table 1 are available in the [Appendix](#).

Section Name	Segment	No. of Question-Items
Demographic details	Demographic details	9
Reviewers' experience with citation of prior literature in manuscripts during manuscripts review	Instances of inadequate and omitted citations in journal manuscripts	4
	Instances of inadequate and omitted citations in conference manuscripts	4
	Factors affecting authors' citing behaviour of prior literature in manuscripts	5

Authors' experience with citation of prior literature in manuscripts during manuscripts review	Instances of inadequate and omitted citations in journal manuscripts	4
	Instances of inadequate and omitted citations in conference manuscripts	4
Researchers' tasks in literature review and manuscript writing process	Tasks where external assistance is required	5
Usage of academic information sources	Usage of academic databases, search engines, digital libraries and scientific paper recommendation services	14

**Table 1: Sections and the corresponding segment in the online questionnaires**

In the second and third sections, data about the different instances of inadequate and omitted citations in journal and conference manuscripts were gathered from the reviewer and author perspectives respectively. The instances *(i) Missed citing seminal papers, (ii) Missed citing topically relevant papers, (iii) Insufficient papers in the literature review and (iv) Irrelevant papers in the literature review* were selected from previous studies which were aimed at helping researchers in literature review ([Ekstrand et al., 2010](#); [Hurtado Martín, Schockaert, Cornelis and Naessens, 2013](#); [Mcnee, 2006](#)). The instances frequency was measured using a 5-point ordinal rating scale with the following values: Never (1), Rarely (2), Sometimes (3), Very Often (4) and Always (5). The second section was hidden for participants who identified themselves as authors with no reviewing experience.

In the fourth section, questions were specifically about tasks pertaining to literature review and manuscript writing. The two segments in this section gathered data about the tasks/scenarios where researchers needed external assistance, along with opinions about the characteristics of certain key tasks. The scenarios were *(i) Identifying seminal/important papers that are to be read as a part of the literature review in your research study, (ii) Identifying papers that are topically similar to the papers that you have already read as part of your literature review, (iii) Identifying papers related to your research, from disciplines other than your primary discipline, (iv) Identifying papers for particular placeholders in your manuscript and (v) Identifying papers that must be necessarily cited in your manuscripts*. These scenarios were identified from previous studies ([Bae et al., 2014](#); [Ekstrand et al., 2010](#); [He, Kifer, Pei, Mitra and Giles, 2011](#); [Hurtado Martín et al., 2013](#); [Mcnee, 2006](#)) where recommender and hybrid information systems were used to provide recommendations to researchers.



The response data was measured using a 5-point ordinal rating scale with the following values: Never (1), Rarely (2), Sometimes (3), Very Often (4) and Always (5).

In the fifth section, data about the usage of information sources such as academic databases, search engines and the *related papers* feature in the aforementioned information sources were collected. The related papers feature retrieves a list of similar papers for a particular paper of interest to the user. The academic information sources considered were *Google Scholar*, *Web of Science*, *ScienceDirect*, *Scopus*, *SpringerLink*, *IEEE Xplore* and *PubMed*. These sources were selected based on their popularity and applicability to multiple disciplines. The response data from the segments in this section were measured using a 5-point ordinal rating scale with the following values: Never (1), Rarely (2), Sometimes (3), Very Often (4) and Always (5).

## Participants

As indicated earlier, two groups of participants were invited to take part in the online surveys pertaining to the objectives of this study. The first group comprised of reviewers who had officially reviewed journal or conference papers. The second group comprised of manuscript authors who had published at least one paper in a journal or conference. The first group was provided with the full questionnaire, with questions that were to be answered from the perspectives of both reviewer and author. The second group was provided with a different survey Web link that pointed to a questionnaire with only author-related questions. The data for the study were collected from a single location - Nanyang Technological University, Singapore between November 2014 and January 2015. The authors sought permission from the research heads of fourteen schools in the university, for disseminating the survey invitation email to the academic staff, research staff and the graduate research students of the schools. For schools that did not respond or declined to disseminate the survey participation email, the required details of the staff and students were obtained from the school Websites and Web directories. Prior to the survey advertisement process and data collection, the required permissions were acquired through the Institutional Review Board of the university. Participants were paid with a cash incentive of 10 Singapore dollars for completing the study.

From a population of 1772 potential participants in the university, a total of 207 researchers (response rate of 12%) participated in the surveys with the majority ( $n = 146$  (71%)) answering from both reviewer and author perspectives, since they met the qualification criteria. Schools from the Engineering

and Natural Sciences disciplines had a bigger presence, with 42% and 33% of the total participants respectively (Figure 1), since the headcount of staff and students in the schools for these disciplines are higher than in the Humanities and Business disciplines. Participants were requested to rate their experience levels. Figure 2 provides a chart with two related demographic variables, experience level and current position. There were an equal number of graduate research students ( $n = 82$ ) and research staff ( $n = 82$ ) while academic staff accounted for 43 participants. Participants were required to provide their journal/conference reviewing and writing experience in years. Figure 3 provides a column chart with experience in number of years for three variables. The three variables are journal review experience, conference review experience and manuscript writing experience. More than half of the participants had less than five years' experience for all the three aforementioned variables: journal review experience (52.3%) conference review experience (54%) and manuscript writing experience (57.9%).

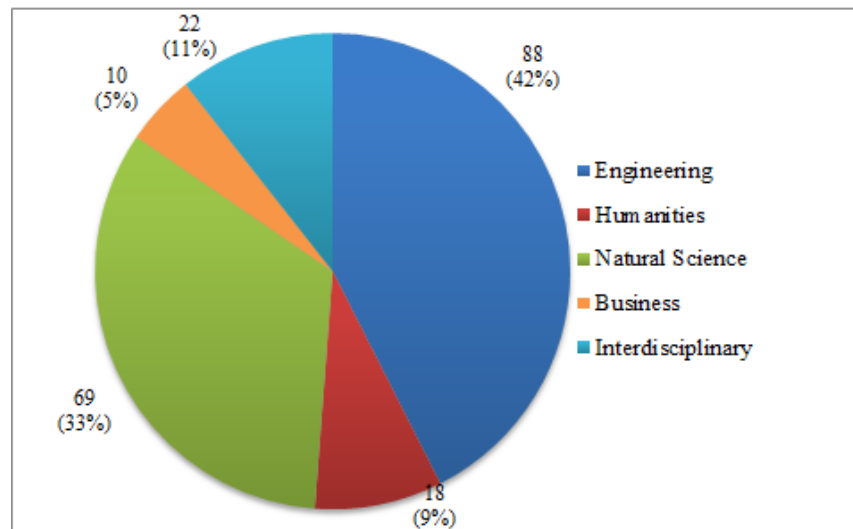
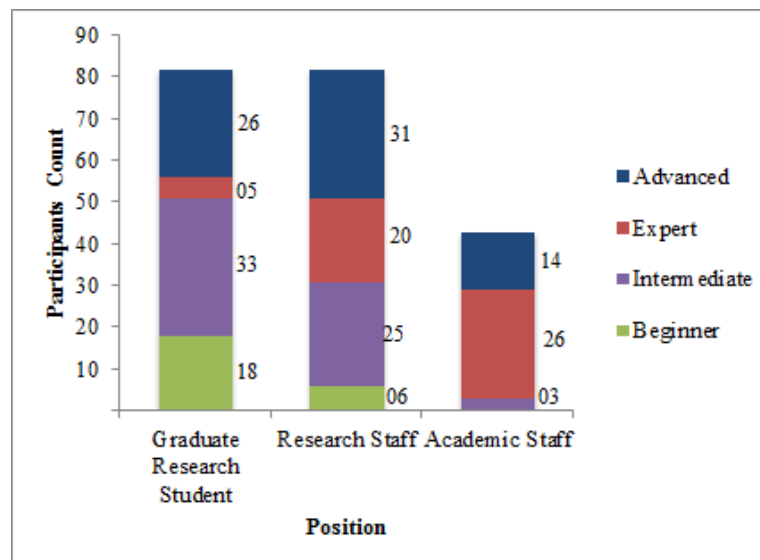
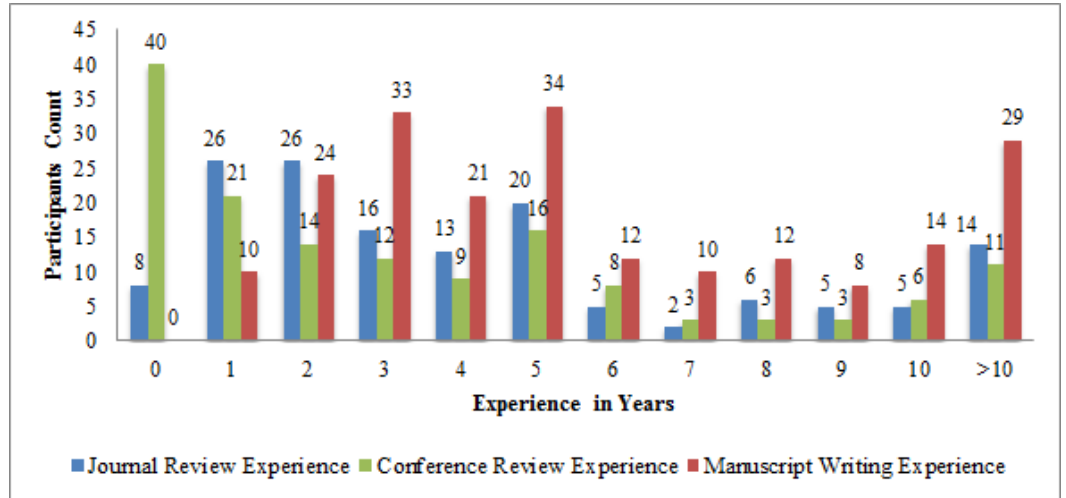


Figure 1: Participants' data by primary discipline



**Figure 2: Participants' data by position and experience level**



**Figure 3: Experience of participants in journal & conference review and writing manuscripts**

## Analysis procedures

Responses to the online survey were collected and analysed. Descriptive statistics were used to measure central tendency. A one-sample t-test was used in the analysis to check the presence of statistically significant difference with the mean values. Statistical significance was set at  $p < 0.05$ . Statistical analyses were done using SPSS 21.0. During the analysis, the continuous variable writing experience was used to create a new categorical variable writing group for facilitating deeper analysis. The writing group 1 (low experience group) was allotted to observations where the participants had indicated their writing experience as less than three years. The writing group 2 (intermediate experience group) was allotted to observations where the writing experience was between three and ten years, while the writing group 3 (high experience group) was allotted to observations where the experience was above 10 years.

## Results

### Instances of inadequate and omitted citations

Table 2 shows the results of the reviewer and author experience responses on the four instances of inadequate and omitted citations for both journal and conference manuscripts. From the reviewer perspective, the mean experience was higher than the test value of 2 (2 represents rare frequency) at statistically significant differences ( $p < 0.05$ ) for all the four instances for journal and conference manuscripts. From the author perspective, only two instances (*Missed citing seminal papers*

*in manuscripts, Missed citing topically relevant papers in manuscripts*) had mean experience higher than the test value of 2, specifically for journal manuscripts. However, all the differences in the author perspective were statistically significant.

The instance *Missed citing topically relevant papers in manuscripts* received the highest mean value for conferences ( $M = 3.14$ ) from a reviewer perspective, followed by the instance *Insufficient papers in the literature review sections of manuscripts* for both journals and conferences ( $M = 3.09$ ,  $M = 3.09$ ). From the author perspectives, results indicate that participants have very infrequently faced these instances in their experience. Only the instances *Missed citing seminal papers in manuscripts* ( $M = 2.29$ ) and *Missed citing topically relevant papers in manuscripts* ( $M = 2.33$ ) are an exception in the case of journal manuscripts.

	Instances of inadequate and omitted citations	Reviewer perspective		Author perspective	
		$M$ (SD)	$t$ ( $p < 0.05$ )	$M$ (SD)	$t$ ( $p < 0.05$ )
Journals	Missed citing seminal papers	2.83 (0.770) <sup>a</sup>	12.718	2.29 (0.888) <sup>c</sup>	4.694
	Missed citing topically relevant papers	3.08 (0.715) <sup>a</sup>	17.729	2.33 (0.865) <sup>c</sup>	5.547
	Insufficient papers in the literature review	3.09 (0.797) <sup>a</sup>	16.027	1.88 (0.881) <sup>c</sup>	-1.972
	Irrelevant papers in the literature review	2.78 (0.896) <sup>a</sup>	10.163	1.41 (0.661) <sup>c</sup>	-12.94
Conferences	Missed citing seminal papers	2.94 (0.780) <sup>b</sup>	12.652	1.79 (0.848) <sup>c</sup>	-3.525
	Missed citing topically relevant papers	3.14 (0.775) <sup>b</sup>	15.317	1.79 (0.859) <sup>c</sup>	-3.479
	Insufficient papers in the literature review	3.09 (0.866) <sup>b</sup>	13.155	1.56 (0.741) <sup>c</sup>	-8.629

Irrelevant papers in the literature review	2.72 (0.901) <sup>b</sup>	8.395	1.32 (0.603) <sup>c</sup>	-16.247
--	------------------------------	-------	------------------------------	---------

**Table 2: Experience of participants on the instances of inadequate and omitted citations**

Sample Size: <sup>a</sup>n = 137, <sup>b</sup>n = 108, <sup>c</sup>n = 207

## Reasons for inadequate and omitted citations

Table 3 shows the results of reviewers' opinions on the reasons for inadequate and omitted citations. Data indicates that there is agreement in support of these reasons. The mean agreement was higher than the test value of 3 (3 represents neutral) at statistically significant differences ( $p < 0.05$ ). The reason *Lack of research experience in particular research area* had the highest average agreement ( $M = 3.68$ ), followed by *Interdisciplinary topic* ( $M = 3.58$ ) and *Lack of overall research experience* ( $M = 3.51$ ).

Reason for inadequate and omitted citations	M (SD)	t ( $p < 0.05$ )
Lack of research experience in particular research area	3.68 (0.813)	10.076
Interdisciplinary topic	3.58 (0.750)	9.382
Lack of overall research experience	3.51 (0.889)	6.982

**Table 3: Participants opinions on reason for inadequate and omitted citations ( $n = 146$ )**

## Need for external assistance in finding papers

Table 4 shows the opinions of the participants on the need for external assistance, for different types of papers, required during literature search sessions in the literature review and manuscript writing lifecycle. The mean value was higher than the test value of 2 (2 represents rare frequency) at statistically significant differences ( $p < 0.05$ ) for the five scenarios. *Interdisciplinary papers* are the paper-type where most external assistance is required during literature search ( $M = 2.72$ ). The need of assistance in searching two other paper-types *Topically-similar papers* ( $M = 2.6$ ) and *Seminal papers* ( $M = 2.53$ ) is also visibly apparent.

Scenarios	M (SD)	t ( $p < 0.05$ )
Seminal papers	2.53 (0.984)	7.696

Typically-similar papers	2.6 (1.009)	8.542
Interdisciplinary papers	2.72 (0.89)	11.71
Citations for placeholders	2.53 (0.954)	7.939
Necessary citations for inclusion in manuscripts	2.48 (1.088)	6.327

**Table 4: Scenarios for external assistance during literature review & academic writing  $n = 207$**

## Usage of academic information sources

Table 5 provides the results of one-sample t-tests carried out with the variables information sources and related papers feature. Test values of 3 (Sometimes) and 2 (Rarely) were used for the two t-tests respectively. *Google Scholar* was the only system with a mean value above the test value ( $M = 3.81$ ) followed by *ScienceDirect* ( $M = 2.86$ ) (not statistically significant at  $p < 0.05$ ) and *Web of Science* ( $M = 2.79$ ). *PubMed* ( $M = 1.9$ ) was the least used of all the information sources. For the variable related papers feature, all the information sources had mean values above the test value and the differences were statistically significant at  $p < 0.05$ . The highest usage was recorded for ScienceDirect's *Recommended Articles* ( $M = 3.11$ ) and Google Scholar's *Related Articles* ( $M = 2.95$ ) while the lowest usage was for SpringerLink's *Related Content* ( $M = 2.4$ ).

Information Source	$M$ (SD)	$t$ ( $p < 0.05$ )	Related Papers Feature	$n$	$M$ (SD)	$t$ ( $p < 0.05$ )
Google Scholar	3.81 (1.137)	10.209	Related Articles	192	2.95 (1.152)	11.402
Web of Science	2.79 (1.374)	-2.227	View Related Records	152	2.6 (1.175)	6.281
Scopus	2.19 (1.373)	-8.455	Related Documents	108	2.83 (1.132)	7.654
IEEE Xplore	2.08 (1.352)	-9.816	Similar	95	2.26 (1.132)	2.266
ScienceDirect	2.86 (1.374)	-1.467	Recommended Articles	151	3.11 (1.105)	12.375
SpringerLink	2.44 (1.241)	-6.442	Related Content	139	2.4 (1.054)	4.505
PubMed	1.9 (1.25)	-12.618	Related Citations	88	2.6 (1.255)	4.501

**Table 5: Usage of academic information sources and related papers feature**

Figure 4 provides the usage response data of information sources by discipline. The respondents who selected any choice other than 1 (Never) in the five-point ordinal Likert scale, were considered for this analysis. Findings indicated that the proportion of usage was similar for sources that indexed research papers from multiple disciplines. Business researchers'

use of Scopus ( $n = 2$ ), IEEE Xplore ( $n = 2$ ) and PubMed ( $n = 4$ ) was minimal, while Humanities researchers' use of IEEE Xplore was almost non-existent ( $n = 1$ ). The highest usage of the sources was from the Engineering respondents with the exception of PubMed where Natural Science researchers' use was the highest ( $n = 35$ ) in comparison to Engineering ( $n = 29$ ).

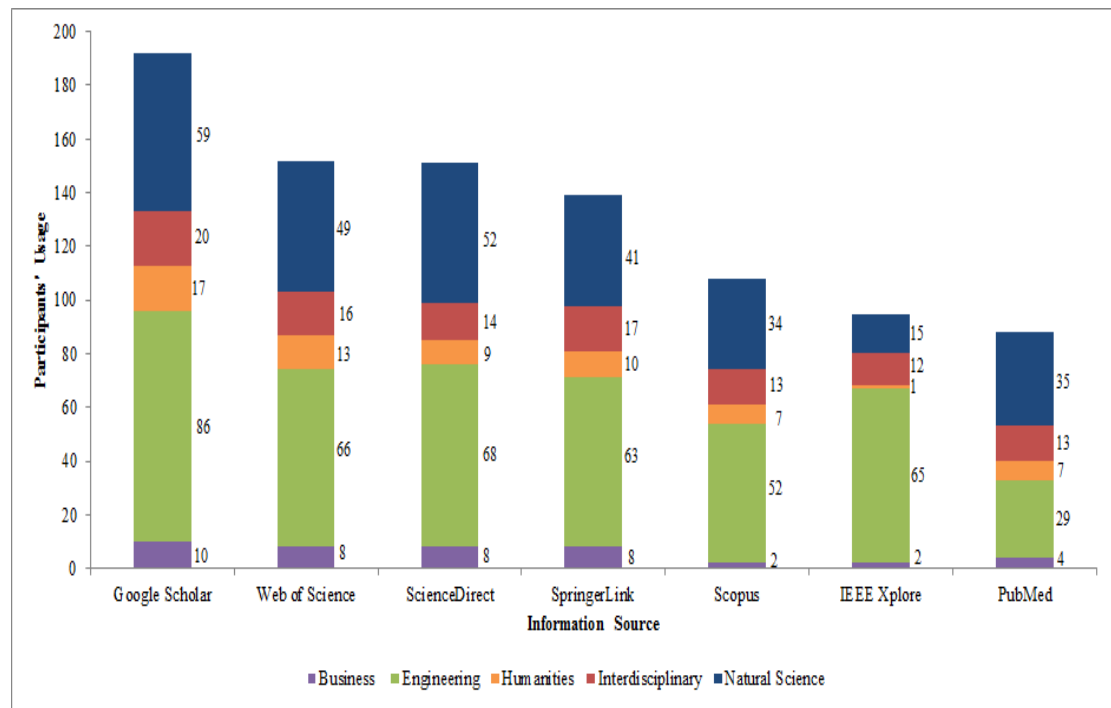


Figure 4: Usage of information sources by discipline

## Discussion

With the exception of the two inadequate citation instances *seminal* and *topically-similar* papers for journal manuscripts, where the perceptions from the two groups were similar, the difference between the reviewers and authors experience is an interesting case. Reviewers have indicated more than a periodic occurrence of the different instances of inadequate and omitted citations, in contrast to the authors' experience. In order to identify reasons behind the low frequency among authors, the data was analysed at the writing group level. The mean values from Table 6 indicate that authors from the medium (group 2) and high experience (group 3) writing groups have experienced these instances more than the authors with low experience (group 1), for the first three key instances. This behaviour is assumed to be due to the relatively low number of papers written by new researchers (doctoral students) and most of these papers tend to be co-written with experts (senior researchers or supervisors) (Heath, 2002). Therefore, authors might probably face these issues as they write more papers in different topics. The presence of the different instances of

inadequate and omitted citations is stronger for journal papers than conference papers. This observation makes sense as conferences are used by researchers mainly for reporting interim results of their research work, in order to acquire suggestive feedback from the research community (Derntl, 2014). In summary for inadequate and omitted citation instances, seminal and topically-similar papers are often missed by authors in their manuscripts, thereby addressing Q1.

Instances of inadequate and omitted citations		Writing Group		
		1 (n = 67) M (SD)	2 (n = 97) M (SD)	3 (n = 43) M (SD)
Journals	Missed citing seminal papers	1.93 (0.893)	2.47 (0.879)	2.44 (0.734)
	Missed citing topically relevant papers	1.96 (0.878)	2.58 (0.814)	2.37 (0.757)
	Insufficient papers in the literature review	1.67 (0.86)	1.95 (0.906)	2.05 (0.815)
	Irrelevant papers in the literature review	1.43 (0.722)	1.38 (0.653)	1.42 (0.587)
Conferences	Missed citing seminal papers	1.72 (0.831)	1.81 (0.894)	1.86 (0.774)
	Missed citing topically relevant papers	1.7 (0.817)	1.87 (0.942)	1.77 (0.718)
	Insufficient papers in the literature review	1.57 (0.763)	1.53 (0.751)	1.6 (0.695)
	Irrelevant papers in the literature review	1.36 (0.644)	1.28 (0.608)	1.35 (0.529)

**Table 6: Mean experience of author writing groups for the instances of inadequate and omitted citations**

When the reasons for inadequate and omitted citations were identified from the literature on graduate information seeking, it was expected that the participants would agree with them as valid reasons. The findings vindicated the expectations as reviewers have opined that the three reasons: *interdisciplinary topic*, *lack of overall research experience* and *lack of research experience in the particular research area* are valid.

Interdisciplinary research presents a unique kind of challenge to the researchers in terms of the information source as there are different academic databases for certain disciplines (George et al., 2006). Even though popular academic search engines such as Google Scholar have a vast coverage of papers from most disciplines, researchers might miss out certain papers which are either not indexed or which do not appear in the top search results (Giustini and Boulos, 2013). General research experience is one of the key factors that differentiate researchers while doing information seeking (Karlsson et al., 2012). However, participants have indicated that the lack of research experience in the particular research area is a more prominent reason than the lack of overall research experience as there are differences in



conducting research across different disciplines.

Researchers have been in situations where they periodically require assistance for finding papers for all the specified tasks or scenarios. For specific search scenarios *finding papers from other disciplines*' and *finding topically-similar papers*, researchers needed more assistance. These search tasks are complex. Interdisciplinary research is a challenging area as it integrates methods from different disciplines ([Wagner et al., 2011](#)). Therefore, researchers would often require assistance in finding papers from other disciplines. The task of finding similar papers is an interesting one as it deals with the aspects of *scale* and *technique*. The *scale* aspect of this task is the quantity of papers that go into the seed set for which similar papers are found. There are various scenarios for this aspect. For example, a researcher might want to find similar papers for one seed paper, while in another scenario the researcher might want to find similar papers for a set of papers ([Mcnee, 2006](#)). The *technique* aspect is about the method and data used in finding similar papers. Some of the commonly used techniques include (i) Citation chaining (forward and backward chaining) based on a set of references ([Levy and Ellis, 2006](#)), (ii) Textual-similarity techniques based on text extracted from a research paper and (iii) Metadata based techniques where metadata from a paper is used to find similar papers. Researchers have indicated that keeping up-to-date with the latest research studies and exploring tangentially similar yet unfamiliar areas are complex tasks ([Athukorala, Hoggan, Lehtio, Ruotsalo and Jacucci, 2013](#)). Therefore, the task is complex where researchers are in situations where they require additional assistance for task completion. Interestingly, the observations for the task of finding similar papers, directly relate to the observations for the corresponding inadequate citation instance. Hence, this task is highly critical for researchers. Secondly, the observations for the task of finding interdisciplinary papers directly relate to the corresponding second most important reason for inadequate and omitted citations. Therefore for Q2, the findings from the three related sections related to each other at a moderate level.

The popularity of Google Scholar among researchers across different disciplines is reinforced through the findings of this study. The finding is an indication of the constant growth in popularity of Google Scholar through the years ([Cothran, 2011](#); [Herrera, 2011](#)). For Q3, the databases ScienceDirect and Web of Science are the other popular choices since they are two of the oldest available online academic information sources, covering multiple disciplines. Most of the databases that were included in the questionnaire are multidisciplinary, with the exception of IEEE Xplore and PubMed. Therefore, these two systems were the least popular sources. The use of all these sources is not

mutually exclusive. The participants' indication of frequently using the related papers feature in the systems suggests their usability and the researcher's dependability on system-level intelligence for the task of finding similar papers. The usage of this feature is highest in the case of ScienceDirect and Google Scholar. In the case of ScienceDirect, the recommended articles are displayed to the right of the currently viewed article. This placement could probably attract more clicks from the users as they get to see topically similar articles placed side-by-side. The usage statistics of these features indicates the reliance of researchers on the system's capability to find similar papers based on multiple factors such as topical similarity, collaborative viewership, shared citations and metadata. Google Scholar's simple and minimalist design requires the users to scroll less and search for options. The interface can be termed as intuitive and sufficient for finding relevant information.

### Implications for assistive systems

The medium of external assistance to researchers can range from online and offline training, to on-site assistance by experts, and also through deployment of assistive intelligent systems. Process-based interventions include strengthening the role of librarians, which is constrained by many factors such as the required knowledge levels and cultural issues ([Ishimura and Bartlett, 2014](#)). On the other hand, academic search systems could be enhanced with task-based features where search results are tailored to the specific search tasks of the researchers. Based on the findings, it is evident that researchers perceive the task of finding similar papers as a manually complex task. Secondly, seminal papers are also missed by manuscript authors, thereby underlining the need for a corresponding recommendation task. These two tasks are intended to address literature review related search activities. Interdisciplinary papers pose a challenge to researchers as they require additional knowledge of terminologies from multiple disciplines. The retrieval/recommendation technique for the two shortlisted tasks should consider interdisciplinary papers as indicated by the explicit preference of the participants in the current study.

The selection of the third task for the assistive system is constrained by the decision to have a task meant to help researchers during manuscript preparation. In this area, citation context recommender system studies ([He et al., 2011](#); [He, Pei, Kifer, Mitra and Giles, 2010](#)) have concentrated on approaches to recommend papers to particular placeholders in the manuscript. Some of the participants in the current study have indicated the issue of insufficient citations in their manuscripts. Therefore, assistive systems could help

researchers in identifying unique and important papers in their final reading list (a list comprising of all papers read or collected during the literature review), just before they start writing their manuscripts. This novel task is taken as the third task of the assistive system. The three selected tasks represent the usual flow of activity in the scientific publication lifecycle. The tasks are logically inter-connected as the papers from the first task become the input to the second and third task. Therefore, a paper-collection feature such as a *seed basket of papers* can be introduced to help the researchers manage the papers between the tasks.

The universal popularity of Google.com as a top general-purpose search engine seems to have been repeated with Google Scholar since most participants, irrespective of discipline, use it frequently. This observation is validated with other studies as well ([Spezi, 2016](#); [Wu and Chen, 2014](#)). However, Google Scholar is classified as a search engine that indexes papers from different academic databases and also from non-academic Websites, while systems such as Scopus and Web of Knowledge are fully-fledged databases with traditional features for search and filtering. In comparison, Google Scholar provides a basic set of advanced search options along with two sorting options. Google Scholar combines a simplistic user-interface with an effective retrieval algorithm to provide fast and relevant results. If the recommendation results in the proposed system are provided in a similar user-interface as Google Scholar, it would benefit the user in quickly adjusting to the new system. Alternatively, if a new user-interface is provided to the users, separate tests are to be conducted for ascertaining the cognitive load and user convenience levels. Therefore, using a user-interface similar to Google Scholar is a recommended approach for the design of assistive systems in this domain. However, the display features of the system need to be tailored as per the nature of the task ([Diriye, Blandford, Tombros and Vakkari, 2013](#)). For instance, a recommendation task for finding similar papers should have a screen where the relations between the recommended papers and the seed papers are to be displayed.

## Limitations

There are two limitations in this study. First, the data have been collected from researchers of a single university; therefore the findings may require validation across other locations. Secondly, certain fundamental differences were perceived between disciplines in the context of research dissemination. For instance, some participants from Natural Science schools indicated that they submit manuscripts mainly to journals while their submissions to conferences were limited. This leads to a scenario where a participant has higher experience in both

authoring and reviewing journal papers than conference papers. At the other end, researchers in Engineering schools submit regularly to both journals and conferences with corresponding review experience levels. Therefore, the survey responses could have been affected by such differences.

## Conclusion

The activities of literature review, execution of research and manuscript writing are the key activities for researchers ([Hart, 1998](#)). Even though there are well-defined guidelines and heuristics given to researchers, the complexity of these activities present a challenge to researchers, particularly those with low experience. The presence of issues with research becomes apparent during the review of research manuscripts. The majority of the manuscript rejections are due to deficiencies in research contributions and methodology ([Bornmann \*et al.\*, 2008, 2009](#); [Byrne, 2000](#)). It is hypothesised that these deficiencies are in-turn caused by an inadequate literature review. Earlier studies have looked at the information seeking patterns of researchers across different disciplines and experience levels, to understand the problems faced during literature search sessions. The perspectives explored so far are that of researchers searching for specific information needs. There is a necessity to learn about the experience and opinions of researchers from the perspectives of manuscript reviewers and authors with the central issue being inadequate and omitted citations in manuscripts.

This paper reports the findings of a survey-based study conducted in a university with participants invited from incumbent schools. 207 researchers with official manuscript review experience and manuscript authoring experience took part in the online survey. Topics in the survey addressed the issue of inadequate and omitted citations in manuscripts. Seminal and topically-similar papers were found to be the most critical instances of citations often missed by authors. Lack of experience in the particular research area was found to be the major reason for missing citations. Participants specifically indicated that they needed external assistance most while finding interdisciplinary papers, seminal papers and topically-similar papers. Based on the findings of the study two literature review search tasks, of finding seminal and topically similar papers, have been identified as the first two tasks of an assistive system. This assistive system is meant to aid researchers in literature search and manuscript writing related tasks. A third task meant to help researchers in identifying unique and important papers from their final reading list, has also been selected for the assistive system. Since Google Scholar was observed to be the predominantly used information source for

searching papers, the user-interface aspects of the search engine have been taken as a basis for designing the user-interface of the assistive system in a similar fashion. The resemblance in the interface should help researchers adjust to the new system in an expedited manner. The findings from the study will help researchers gain a rich understanding of the issue of inadequate and omitted citations in manuscripts since data have been collected from the dual perspectives of manuscript reviewers and authors.

## About the authors

**Aravind Sesagiri Raamkumar** is a PhD Candidate in the Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore. He received his MSc in Knowledge Management from Nanyang Technological University. His research interests include recommender systems, information retrieval, scholarly communication, social media and linked data. He can be contacted at [aravind002@ntu.edu.sg](mailto:aravind002@ntu.edu.sg).

**Schubert Foo** is Professor of Information Science at the Wee Kim Wee School of Communication and Information, Nanyang Technological University (NTU), Singapore. He received his B.Sc. (Hons), M.B.A. and Ph.D. from the University of Strathclyde, UK. He has published more than 300 publications in his research areas of multimedia and Internet technologies, multilingual information retrieval, digital libraries, information and knowledge management, information literacy and social media innovations. He can be contacted at [sfoo@ntu.edu.sg](mailto:sfoo@ntu.edu.sg).

**Natalie Pang** is an Assistant Professor in the Wee Kim Wee School of Communication and Information, Nanyang Technological University. Her research interest is in the area of social informatics, focusing on basic and applied research of social media, information behaviour in crises, and structural models of technology use in marginalised communities. She can be contacted at [nlspang@ntu.edu.sg](mailto:nlspang@ntu.edu.sg).

## References

- Atanassova, I. & Bertin, M. (2014). Faceted semantic search for scientific papers. *PLoS Biology*, 2, 426-522.
- Athukorala, K., Hoggan, E., Lehtio, A., Ruotsalo, T. & Jacucci, G. (2013). Information-seeking behaviors of computer scientists: challenges for electronic literature search tools. *Proceedings of the American Society for Information Science and Technology*, 50(1), 1-11.
- Bae, D-H., Hwang, S-M., Kim, S-W. & Faloutsos, C. (2014). On constructing seminal paper genealogy. *IEEE Transactions on Cybernetics*, 44(1), 54-65.
- Barrett, A. (2005). The information-seeking habits of graduate student researchers in the humanities. *The Journal of*

- Academic Librarianship*, 31(4), 324–331.
- Bates, M. J. (2002). Toward an integrated model of information seeking and searching. *New Review of Information Behaviour Research*, 3, 1–15.
- Beel, J., Langer, S., Genzmehr, M. & Nürnberger, A. (2013). Introducing Docear's research paper recommender system. In *Proceedings of the 13th ACM/IEEE-CS joint conference on Digital libraries* (pp. 459-460). New York, NY: ACM Press.
- Björk, B.-C. & Hedlund, T. (2003). [\*Scientific publication life-cycle model \(SPLC\)\*](#). In Sely Maria de Souza Costa, João Álvaro Carvalho, Ana Alice Baptista and Ana Cristina Santos Moreira, (Eds). *ELPUB2003. From information to knowledge: Proceedings of the 7th ICC/IFIP International Conference on Electronic Publishing held at the Universidade do Minho, Portugal 25-28 June 2003*. Retrieved from <http://elpub.scix.net/data/works/att/0317.content.pdf> (Archived by WebCite® at <http://www.webcitation.org/6llMkyD7z>)
- Bordage, G. (2001). Reasons reviewers reject and accept manuscripts: the strengths and weaknesses in medical education reports. *Academic Medicine*, 76(9), 889–896.
- Bornmann, L., Nast, I. & Daniel, H-D. (2008). Do editors and referees look for signs of scientific misconduct when reviewing manuscripts? A quantitative content analysis of studies that examined review criteria and reasons for accepting and rejecting manuscripts for publication. *Scientometrics*, 77(3), 415–432.
- Bornmann, L., Weymuth, C. & Daniel, H-D. (2009). A content analysis of referees' comments: how do comments on manuscripts rejected by a high-impact journal and later published in either a low- or high-impact journal differ? *Scientometrics*, 83(2), 493–506.
- Brand-Gruwel, S., Wopereis, I. & Vermetten, Y. (2005). Information problem solving by experts and novices: analysis of a complex cognitive skill. *Computers in Human Behavior*, 21(3), 487–508.
- Braun, T. (2004). Keeping the gates of science journals. In *Handbook of quantitative science and technology research* (pp. 95–114). Dordrecht, The Netherlands: Springer Netherlands.
- Byrne, D. W. (2000). Common reasons for rejecting manuscripts at medical journals: a survey of editors and peer reviewers. *Science Editor*, 23(2), 39–44.
- Case, D. O. & Higgins, G. M. (2000). How can we investigate citation behavior? A study of reasons for citing literature in communication. *Journal of the American Society for Information Science*, 51(7), 635–645.
- Case, D. O. & Miller, J. B. (2011). Do bibliometricians cite differently from other scholars? *Journal of the American Society for Information Science and Technology*, 62(3), 421–432.
- Catalano, A. (2013). Patterns of graduate students' information

- seeking behavior: a meta-synthesis of the literature. *Journal Of Documentation*, 9(2), 243–274.
- Chou, J-K. & Yang, C-K. (2011). PaperVis: literature review made easy. *Computer Graphics Forum*, 30(3), 721–730.
- Cothran, T. (2011). Google Scholar acceptance and use among graduate students: a quantitative study. *Library & Information Science Research*, 33(4), 293–301.
- Derntl, M. (2014). Basics of research paper writing and publishing. *International Journal of Technology Enhanced Learning*, 6(2), 105–123.
- Diriye, A., Blandford, A., Tombros, A. & Vakkari, P. (2013). The role of search interface features during information seeking. In *Research and advanced technology for digital libraries* (pp. 235–240). Berlin, Heidelberg: Springer.
- Du, J. T. & Evans, N. (2011). Academic users' information searching on research topics: characteristics of research tasks and search strategies. *The Journal of Academic Librarianship*, 37(4), 299–306.
- Ekstrand, M. D., Kannan, P., Stemper, J. A., Butler, J. T., Konstan, J. A. & Riedl, J. T. (2010). Automatically building research reading lists. In *Proceedings of the Fourth ACM Conference on Recommender Systems* (pp. 159–166). New York, NY: ACM Press.
- Ellis, D. & Haugan, M. (1997). Modelling the information seeking patterns of engineers and research scientists in an industrial environment. *Journal of Documentation*, 53(4), 384–403.
- Erikson, M. G. & Erlandson, P. (2014). A taxonomy of motives to cite. *Social Studies of Science*, 44(4), 625–637.
- Fidzani, B. T. (1998). Information needs and information-seeking behaviour of graduate students at the University of Botswana. *Library Review*, 47(7), 329–340.
- Gamberini, L., Spagnolli, A., Blankertz, B., Kaski, S., Freeman, J., Acqualagna, L., ... & Jacucci, G. (2015). Developing a symbiotic system for scientific information seeking: the MindSee Project. In B. Blankertz, G. Jacucci, L. Gamberini, A. Spagnolli, & J. Freeman (Eds.), *Symbiotic Interaction: 4th International Workshop, Symbiotic 2015: proceedings* (pp. 68–80). Cham, Switzerland: Springer International Publishing. (Lecture Notes in Computer Science, Vol, 9359)
- George, C. A., Bright, A., Hurlbert, T., Linke, E. C., Clair, G. S. & Stein, J. (2006). [Scholarly use of information: graduate students' information seeking behaviour](#). *Information Research*, 11(4), paper 272. Retrieved from <http://www.informationr.net/ir/11-4/paper272.html> (Archived by WebCite® at <http://www.webcitation.org/6llNuGXl8>)
- Giustini, D. & Boulos, M. N. K. (2013). [Google Scholar is not enough to be used alone for systematic reviews](#). *Online Journal of Public Health Informatics*, 5(2). Retrieved from <http://journals.uic.edu/ojs/index.php/ojphi/article/view/4623> (Archived by WebCite® at <http://www.webcitation.org/6llPs5oQW>)
- Hart, C. (1998). *Doing a literature review: releasing the social*

- science research imagination*. London: Sage Publications.
- Harwood, N. (2009). An interview-based study of the functions of citations in academic writing across two disciplines. *Journal of Pragmatics*, 41(3), 497–518.
- He, Q., Kifer, D., Pei, J., Mitra, P. & Giles, C. L. (2011). Citation recommendation without author supervision. In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining - WSDM '11* (p. 755). New York, NY: ACM Press.
- He, Q., Pei, J., Kifer, D., Mitra, P. & Giles, L. (2010). Context-aware citation recommendation. In *Proceedings of the 19th International Conference on World Wide Web - WWW '10* (pp. 421-430). New York, NY: ACM Press.
- [Head, A. J. \(2007\). Beyond Google: how do students conduct academic research? \*First Monday\*, 12\(8\). Retrieved from http://ojs-prod-lib.cc.uic.edu/ojs/index.php/fm/article/view/1998](http://ojs-prod-lib.cc.uic.edu/ojs/index.php/fm/article/view/1998)  
(Archived by WebCite® at <http://www.webcitation.org/6llQXnRV>)
- Heath, T. (2002). A quantitative analysis of PhD students' views of supervision. *Higher Education Research & Development*, 21(1), 41–53.
- Herrera, G. (2011). Google Scholar users and user behaviors: an exploratory study. *College & Research Libraries*, 72(4), 316–330.
- Huang, W., Wu, Z., Mitra, P. & Giles, C. L. (2014). RefSeer : a citation recommendation system. In *Proceedings of the 14th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 371–374). Piscataway, NJ: IEEE Press.
- Hurtado Martín, G., Schockaert, S., Cornelis, C. & Naessens, H. (2013). Using semi-structured data for assessing research paper similarity. *Information Sciences*, 221, 245–261.
- Ishimura, Y. & Bartlett, J. C. (2014). Are librarians equipped to teach international students? A survey of current practices and recommendations for training. *Journal of Academic Librarianship*, 40(3), 313–321.
- Karlsson, L., Koivula, L., Ruokonen, I., Kajaani, P., Antikainen, L. & Ruismäki, H. (2012). From novice to expert: information seeking processes of university students and researchers. *Procedia - Social and Behavioral Sciences*, 45, 577–587.
- Küçüktunç, O., Saule, E., Kaya, K. & Çatalyürek, Ü. V. (2013). TheAdvisor : a Webservice for academic recommendation. In *Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 433–434). New York, NY: ACM Press.
- Levy, Y. & Ellis, T.J. (2006). A systems approach to conduct an effective literature review in support of information systems research. *Informing Science: International Journal of an Emerging Transdiscipline*, 9(1), 181–212.
- Liu, J. & Kim, C. S. (2013). Why do users perceive search tasks as difficult? Exploring difficulty in different task types. In *Proceedings of the Symposium on Human-Computer*



- Interaction and Information Retrieval* (pp. 1–10). New York, NY: ACM Press.
- MacRoberts, M. H. & MacRoberts, B. R. (1988). Author motivation for not citing influences: a methodological note. *Journal of the American Society for Information Science*, 39(6), 432–433.
- MacRoberts, M. H. & MacRoberts, B. R. (1989). Problems of citation analysis: a critical review. *Journal of the American Society for Information Science*, 40(5), 342–349.
- MacRoberts, M. H. & MacRoberts, B. R. (2010). Problems of citation analysis: a study of uncited and seldom-cited influences. *Journal of the American Society for Information Science and Technology*, 61(1), 1–13.
- McKercher, B., Law, R., Weber, K. & Hsu, C. (2007). Why referees reject manuscripts. *Journal of Hospitality & Tourism Research*, 31(4), 455–470.
- McNee, S. M. (2006). *Meeting user information needs in recommender systems*. (Unpublished doctoral dissertation). University of Minnesota, Minneapolis, Minnesota, USA.
- Murthy, U. S. & Wiggins, C. E. (2002). Why manuscripts are rejected: an analysis of JIS rejections. *Journal of Information Systems*, 16(1), 41–48.
- Naak, A., Hage, H. & Aïmeur, E. (2008). Papyrus: a research paper management system. In *10th IEEE Conference on E-Commerce Technology and the Fifth IEEE Conference on Enterprise Computing, E-Commerce and E-Services* (pp. 201–208). Piscataway, NJ: IEEE Press.
- Niu, X. & Hemminger, B. M. (2012). A study of factors that affect the information-seeking behavior of academic scientists. *Journal of the American Society for Information Science and Technology*, 63(2), 336–353.
- Oppenheim, C. & Renn, S. P. (1975). Highly cited old papers and the reasons why they continue to be cited. *Journal of the American Society for Information Science*, 29(5), 225–231.
- Simkin, M. V. & Roychowdhury, V. P. (2002). [Read before you cite!](#) *Complex Systems*, 14, 269–274. Retrieved from [http://www.complex-systems.com/abstracts/v14\\_i03\\_a05.html](http://www.complex-systems.com/abstracts/v14_i03_a05.html) (Archived by WebCite® at <http://www.webcitation.org/6llRuSP19>)
- Spezi, V. (2016). Is information-seeking behavior of doctoral students changing? A review of the literature (2010–2015). *New Review of Academic Librarianship*, 2(1), 78–106.
- Sturm, B., Schneider, S. & Sunyaev, A. (2015). [Leave no stone unturned: introducing a revolutionary meta-search tool for rigorous and efficient systematic literature searches](#). In *Proceedings of 23rd European Conference on Information Systems (ECIS 2015)*. Retrieved from [http://aisel.aisnet.org/ecis2015\\_rip/34/](http://aisel.aisnet.org/ecis2015_rip/34/) (Archived by WebCite® at <http://www.webcitation.org/6llSaOWwK>)
- Vilar, P. & Žumer, M. (2011). Information searching behaviour of young Slovenian researchers. *Program*, 45(3), 279–293.
- Wagner, C. S., Roessner, J. D., Bobb, K., Klein, J. T., Boyack, K.

- W., Keyton, J., ... Börner, K. (2011). Approaches to understanding and measuring interdisciplinary scientific research (IDR): a review of the literature. *Journal of Informetrics*, 5(1), 14–26.
- Wang, Y., Zhai, E., Hu, J. & Chen, Z. (2010). Claper: recommend classical papers to beginners. In *2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2010)* (pp. 2777–2781). Piscataway, NJ: IEEE Press.
- Wildemuth, B. M. & Freund, L. (2012). Assigning search tasks designed to elicit exploratory search behaviors. In *Proceedings of the Symposium on Human-Computer Interaction and Information Retrieval - HCIR '12* (pp. 1–10). New York, NY: ACM Press.
- Willett, P. (2013). Readers' perceptions of authors' citation behaviour. *Journal of Documentation*, 69(1), 145–156.
- Wu, M. & Chen, S. (2014). Graduate students appreciate Google Scholar, but still find use for libraries. *The Electronic Library*, 32(3), 375–389.
- Yoo, I. & Mosa, A. S. M. (2015). Analysis of PubMed user sessions using a full-day PubMed query log: a comparison of experienced and nonexperienced PubMed users. *JMIR Medical Informatics*, 3(3), e25.

#### How to cite this paper

Raamkumar, A.S., Foo, S. & Pang, N. (2016). Survey on inadequate and omitted citations in manuscripts: a precursory study in identification of tasks for a literature review and manuscript writing assistive system. *Information Research*, 21(4), paper 733. Retrieved from <http://InformationR.net/ir/21-4/paper733.html> (Archived by WebCite® at <http://www.webcitation.org/6m5HZifOk>)

Find other papers on this subject

Check for citations, [using Google Scholar](#)

Facebook

Twitter

LinkedIn

Delicious

More

## Appendix

### Survey questionnaire

#### Section 1: Demographics Details

1. Please provide your initials or name

2. Gender [Male/Female]
3. Age [18-24 years old/25-34 years old/35-44 years old/45-54 years old/55-64 years old/65 years or older]
4. Highest Education Level [Post Graduate Diploma/Bachelor's Degree/ Master's Degree/Doctor's Degree/Others]
5. Current Position [Graduate Research Student/Research Associate/Research Fellow/Lecturer/Senior Lecturer/Assistant Professor/Associate Professor/Professor/Others]
6. Parent School
7. Primary Discipline
8. Have you played the role of a reviewer for journals or conferences in the past? [Yes/No]
9. Which of the below research experience level describes you the best? [Expert/Advanced/Intermediate/Beginner]

## **Section 2: Reviewers' experience with citation of prior literature in manuscripts during manuscript's review**

10. For how many years, have you been reviewing journal and conference papers?
11. While reviewing journal paper manuscripts, I have come across instances where...  
[Never/Rarely/Sometimes/Very Often/Always]
  - a. Authors have missed citing important/seminal papers in their manuscripts
  - b. Authors have missed citing topically relevant papers in their manuscripts
  - c. There was need for authors to cite more papers in the Literature Review/Related Work sections of their manuscripts
  - d. Authors have cited irrelevant papers in their manuscripts
12. While reviewing conference paper manuscripts, I have come across instances where...  
[Never/Rarely/Sometimes/Very Often/Always]
  - a. Authors have missed citing important/seminal papers in their manuscripts
  - b. Authors have missed citing topically relevant papers in their manuscripts
  - c. There was need for authors to cite more papers in the Literature Review/Related Work sections of their manuscripts
  - d. Authors have cited irrelevant papers in their manuscripts
13. Please provide your opinion on factors affecting authors' citing behavior of prior literature in manuscripts  
[Strongly Disagree/Disagree/Neutral/Agree/Strongly Agree]
  - a. Authors miss citing relevant papers particularly

- when the research topic is interdisciplinary in nature
- b. Authors' tendency of not citing the expected papers in their manuscripts, can be attributed to the lack of overall research experience
  - c. Authors' tendency of not citing the expected papers in their manuscripts, can be attributed to the lack of experience in the particular research area
  - d. Authors' inability in citing the relevant literature is one of the key reasons for sending the manuscript back for revision
  - e. Authors' inability in citing the relevant literature is one of the key reasons for rejecting the manuscript

### **Section 3: Authors' Experience with Citation of Prior Literature in Manuscripts during Manuscripts Review**

14. For how many years, have you been writing papers for journals and conferences?
15. During submission of journal paper manuscripts, I have come across instances where...  
[Never/Rarely/Sometimes/Very Often/Always]
  - a. Reviewers have indicated that I have not cited certain important/seminal papers in the manuscript
  - b. Reviewers have indicated that I have not cited certain topically related paper(s), in the manuscript
  - c. Reviewers have indicated that I have not cited enough papers in the Literature Review section of the manuscript
  - d. Reviewers have indicated that I have cited irrelevant papers, in the manuscript
16. During submission of conference paper manuscripts, I have come across instances where...  
[Never/Rarely/Sometimes/Very Often/Always]
  - a. Reviewers have indicated that I have not cited certain important/seminal papers in the manuscript
  - b. Reviewers have indicated that I have not cited certain topically related paper(s), in the manuscript
  - c. Reviewers have indicated that I have not cited enough papers in the Literature Review section of the manuscript
  - d. Reviewers have indicated that I have cited irrelevant papers, in the manuscript

### **Section 4: Researchers' Tasks in Literature Review and Manuscript Writing Process**

17. How often do you feel the need for assistance in...  
[Never/Rarely/Sometimes/Very Often/Always]
  - a. Identifying seminal/important papers that are to be

read as a part of the literature review in your research study

- b. Identifying papers that are topically similar to the papers that you have already read as part of your literature review
- c. Identifying papers related to your research, from disciplines other than your primary discipline
- d. Identifying papers for particular placeholders in your manuscript
- e. Identifying papers that must be necessarily cited in your manuscripts

## **Section 5: Usage of academic information sources**

- 18. I have used the academic search engine 'Google Scholar' during my information search sessions  
[Never/Rarely/Sometimes/Very Often/Always]
  - 19. I have used the 'Related Articles' feature in Google Scholar  
[Never/Rarely/Sometimes/Very Often/Always]
  - 20. I have used the academic database 'Web of Science' during my information search sessions  
[Never/Rarely/Sometimes/Very Often/Always]
  - 21. I have used the 'View Related Records' feature in Web of Science [Never/Rarely/Sometimes/Very Often/Always]
  - 22. I have used the academic database 'Scopus' during my information search sessions  
[Never/Rarely/Sometimes/Very Often/Always]
  - 23. I have used the 'Related Documents' feature in Scopus  
[Never/Rarely/Sometimes/Very Often/Always]
  - 24. I have used the academic database 'IEEE Xplore' during my information search sessions  
[Never/Rarely/Sometimes/Very Often/Always]
  - 25. I have used the 'Similar' feature in IEEE Xplore  
[Never/Rarely/Sometimes/Very Often/Always]
  - 26. I have used the academic database 'Science Direct' during my information search sessions  
[Never/Rarely/Sometimes/Very Often/Always]
  - 27. I have used the 'Recommended Articles' feature in Science Direct [Never/Rarely/Sometimes/Very Often/Always]
  - 28. I have used the academic database 'Springer Link' during my information search sessions  
[Never/Rarely/Sometimes/Very Often/Always]
  - 29. I have used the 'Related Content' feature in Springer Link  
[Never/Rarely/Sometimes/Very Often/Always]
  - 30. I have used the academic search engines 'PubMed' during my information search sessions  
[Never/Rarely/Sometimes/Very Often/Always]
  - 31. I have used the 'Related Citations' feature in PubMed  
[Never/Rarely/Sometimes/Very Often/Always]
-

© the authors, 2016.

**56** Last updated: 16 November, 2016

---

[Contents](#) | [Author index](#) | [Subject index](#) | [Search](#) |  
[Home](#)

---