

In their own image? A comparison of doctoral students' and faculty members' referencing behavior

Vincent Larivière^{1,2}, Cassidy R. Sugimoto³ and Pierrette Bergeron¹

¹ École de bibliothéconomie et des sciences de l'information, Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal, QC. H3C 3J7, Canada

² Observatoire des Sciences et des Technologies (OST), Centre Interuniversitaire de Recherche sur la Science et la Technologie (CIRST), Université du Québec à Montréal, CP 8888, Succ. Centre-Ville, Montréal, QC. H3C 3P8, Canada

³ School of Library and Information Science, Indiana University, 1320 East Tenth Street, Bloomington, IN 47405-3907, USA

[vincent.lariviere@umontreal.ca; sugimoto@indiana.edu; pierrette.bergeron@umontreal.ca]

ABSTRACT

This paper compares doctoral students and faculty members' referencing behavior through the analysis of a large corpus of scientific papers. It shows that doctoral students tend to cite more documents per paper than faculty members, and that the literature they cite is, on average, more recent. It also demonstrates that doctoral students cite a larger proportion of conference proceedings and journal articles than faculty members and faculty members are more likely to self-cite and cite theses than doctoral students. Analysis of the impact of cited journals indicates that, in health research, faculty members tend to cite journals with slightly lower impact factors while in SSH, faculty members cite journals with higher impact factors. Finally, it provides evidence that, in every discipline, faculty members tend to cite a higher proportion of clinical/applied research journals than doctoral students. This paper contributes to our understanding of referencing patterns and age stratification in academia. Implications for our understanding of information seeking behavior of academics are discussed.

INTRODUCTION

Two pillars of contemporary information science research are bibliometrics and information seeking behavior (Milojević, Sugimoto, Yan, & Ding, 2011). However, scant research explicitly connects these subfields. We argue that these are inextricably linked. The process by which scholars actively and passively accrue information provides the parameters for the information that they can select for inclusion in a scientific article. Studies on citation motivation provide insight into the rationale of selecting citations and the function references serve in a manuscript. Reference analysis provides insight into patterns of referencing across documents aggregated by author, discipline, country, time or other variable. To these areas might be added doctoral mentoring, as this is arguably where students are inculcated to the norms of the discipline and should learn, implicitly or explicitly, how to write and cite.

It has been noted that ideal mentors are those who reproduce mentees in their own image (Haring, 1999). In academe, the mentoring relationship between advisors and their doctoral students is no different. As "gate keepers to the scholarly profession" and the "socializing agents of the discipline"

(Girves & Wemmerus, 1998, p. 171) one might expect that faculty members seek to reproduce students with similar scholarly norms and practices. This would reinforce a notion of invisible college (de Solla Price, 1966) or disciplinary communities of practices, whereby a group of individuals are “bound by common interests” and “interact through interdependent tasks guided by a common purpose thereby embodying a store of common knowledge” (Jubert [1999, p. 166] as cited by Davenport and Hall [2002]). As Lave and Wenger (1991) describe, participants of a community of practice engage in a process of learning in a social context in which they move through various stages of initiation to become an expert. As noted by Davies, “individuals learn...through the process of becoming a full participant in a sociocultural practice” (Davies, 2005, p. 104).

Two defining practices of scholar are the seeking and use of previous research in the writing of new scholarly works. However, rather than replicating behaviors, we find evidence of variation between doctoral students and faculty members in these activities. For example, in a study of more than 2,000 students, staff, and faculty members in five U.S. institutions, academic position was found to be the most influential determinant of information seeking behavior (Niu & Hemminger, 2012). These variations in information seeking behavior materialize in the form of variations in referencing patterns. Previous literature shows that younger scholars cite younger literature (e.g., Barnett & Fink, 2008; Gingras, Larivière, Macaluso, & Robitaille, 2008; Milojević, 2011; Pancheshnikov, 2007; Zuckerman & Merton, 1973). Barnett and Fink (2008) propose three reasons for this behavior: a) younger scholars are more receptive to new ideas than young scholars; b) older scholars have an older set of literature as their primary knowledge base; and c) older scholars serve as gatekeepers. In addition, doctoral students are exposed to a “mentoring constellation” (Sugimoto, 2012) that may expose them to “a wider variety of fields than their older and temporarily more specialized teachers” (Zuckerman & Merton, 1973, p. 511).

The citing behavior might also be influenced by the type of science and research conducted, as well as by the exponential increase in scientific literature, as argued by Luker (2008). She claims that in what she calls the age of “canonical social sciences”, a high consensus existed in one discipline about who had to be cited, and doctoral students had “to master the oldies but goodies and the up-and-comers” (Luker, 2008, p. 76). However, in the social sciences at least, with the increasingly multidisciplinary approach required to conduct research and tackle complex social phenomena, and the use of qualitative research methods, literature reviews are no longer based on this model of citing the “consensus”, since one has to cover more than one discipline (Luker, 2008). Given the growing number of doctoral students entering scientific professions, the changing age distribution might result in a “juvanescent effect” that could dramatically change the norms of the scientific community (Zuckerman & Merton, 1973, p. 500). However, few studies have examined the relationship between age stratification and referencing on a macro level (Gingras, Larivière, Macaluso, & Robitaille, 2008). Furthermore, few studies address these differences from a domain analytic approach, that is, understanding that information seeking and use vary by disciplinary areas (e.g., Hjørland & Albrechtsen, 1995; Talja & Maula, 2003).

This paper attempts to address this problem by investigation of potential differences in the referencing patterns of all doctoral students and faculty members in Quebec universities for the period 2000 to 2007

looking across all disciplinary areas (i.e., health, natural sciences, and social sciences and humanities). Specifically, this paper investigates the following research questions and sub-questions:

RQ1. Are there significant differences between faculty members and doctoral students in the number of references used? Does this vary by disciplinary area?

RQ2. Are there significant differences between faculty members and doctoral students in the age of references used? Does this vary by disciplinary area?

RQ3. Are there significant differences between faculty members and doctoral students in the type of genres cited? Does this vary by disciplinary area?

RQ4. Are there significant differences between faculty members and doctoral students in the number of self-citations (at the personal and national level)? Does this vary by discipline?

RQ5. Are there significant differences between faculty members and doctoral students in the impact factor of sources cited? Does this vary by discipline?

RQ6. Are there significant differences between faculty members and doctoral students in the research level of journals cited? Does this vary by discipline?

This paper presents a comprehensive analysis of a comparable set of doctoral students and faculty members over an extended time period and across all disciplines represented in the Web of Science. Although earlier studies have addressed some of these questions looking at specific disciplines or institutions, this study attempts to examine from a macro level and from a number of perspectives (e.g., research levels and impact of citations). This comprehensive study contributes to our understanding of age stratification in science and provides insight on the implications of age differences for the trajectory of scholarly communication.

BACKGROUND

This paper is informed and contributes to literature on information seeking behavior, citation motivation, referencing behavior, and doctoral mentoring in regards to information seeking and use. Relevant studies in these areas are briefly discussed below.

Information seeking behavior of academics

There are more studies on the information seeking behavior of academics than any other group (Case, 2007). A large number of studies have looked specifically at the information seeking behavior of graduate students (e.g., Barrett, 2005; Cole, 1997; Cole, 1998; Cole, 2000; Delgadillo & Lynch, 1999; Fleming-May & Yuro, 2009; George, Bright, Hurlbert, Linke, St. Clair, & Stein, 2006; Jamali & Nicholas, 2008; Khosrowjerdi & Iranshahi, 2011; Kim, 2009; Korobili, Malliari, & Zapounidou, 2011; Liao, Finn, & Lu, 2007). Faculty members have also been a targeted population for study (e.g., Gao, Guo, & Duan, 2011; Meho & Haas, 2001; Meho & Tibbo, 2003), with some articles focused on a subset of scholars (e.g., social scientists) (Line, 1971). However, few studies have conducted comparative analyses of

faculty and doctoral students. Jamali and Nicholas (2008) found that doctoral students associated less importance to “keeping up-to-date” than senior researchers. However, this conflicts with research that suggests that younger researchers read more than older academics and use more references (Tenopir, King, Spencer, & Wu, 2009; cited by Frandsen & Nicolaisen, 2012). Doctoral students reportedly use word-of-mouth and browsing e-journals more than senior researchers (Jamali & Nicholas, 2008). Similarly, in interviews with humanities faculty and students, Ge (2010) found that doctoral students and assistant faculty members relied on electronic materials more than associate and full professors. However, most of these studies have been focused on small case studies of single disciplines or single institutions, diminishing their generalizability across the scientific landscape.

Citation motivation

There have been many studies that have reviewed theoretical, conceptual and empirical studies of referencing behavior from the point of view of citation motivation and function (e.g., Bornmann & Daniel, 2008; Brooks, 1986; Case & Higgins, 2000; Cronin, 1981; Hellqvist, 2010; Wang & White, 1999). These studies attempt to address why scholars choose to include certain references and the ways in which the references are used. Few studies make explicit the connection between information seeking and citing. Wang and White (1999) indicate that respondents mentioned more criteria for deciding whether to cite a document than to select or read it—a finding that could be expected given the level of ‘commitment’ associated with citing. There is “a progression in the researcher’s orientation from essentially private use of documents during the research to public use, that is, documenting that use for peers, referees, editors and funding agencies” (Wang & White, 1999, p. 111). The authors found that respondents with a higher degree of expertise in their topic choose fewer documents, read more, and cite more. This suggests that studies of information seeking behavior, scientific aging, and referencing practices might reinforce and inform one another, providing fertile ground for triangulation.

Reference studies

Reference studies attempt to describe patterns across documents by examining similarities in referencing behavior. White (2001) pioneered the egocentric approach, examining the aggregate references of a few select LIS scholars and introduced the concept of citation identities, that is, “the set of all authors whom an author cites”. However, “studies that explore the characteristics of citation practices and of references cited on a large scale (not at the level of an individual, or small group of scientists), and over extended periods of time are rare” (Milojević, 2011, p.1). Milojević’s (2011) study addresses this gap by examining the reference practices of LIS faculty, examining the relationship between age of the references and seniority of researchers, finding that senior authors use fewer and older references. In addition, she noted that high levels of recitations were positively correlated with author productivity.

Many studies describe referencing patterns in specific disciplines. These studies often focus on elements such as genre type, age of cited material, language of cited material, number of references and most cited items in order to provide information about particular disciplines (e.g., Budd, 1990, Creaser, Oppenheim, & Summers, 2011; Silverman, 1985). They suggest that there are variations in production

and referencing practices between fields (e.g., Creaser, Oppenheim, & Summers, 2011; Kidd, 1990; Larivière, Archambault, Gingras, & Vignola-Gagné, 2006; Nederhof, 2011).

Many referencing studies that focus on the products of graduate and doctoral students have been done to inform collection development in the library (e.g., Keat & Kiran, 2008; Kumar & Dora, 2011; Wu & Chen, 2010) and often involve a comparison with the local collection (e.g., Feyereisen & Spoiden, 2009; Sherriff, 2010). The studies of theses examine many of the same variables examined in other referencing studies: age, language, format, discipline of cited materials and well as the most cited authors and venues (e.g., Buttlar, 1999; Conkling, Harwell, McCallips, Nyana, & Osif, 2010; Feyereisen & Spoiden, 2009; Gooden, 2001; Keat & Kiran, 2008; Kumar & Dora, 2011; Kushkowsky, Parsons, & Wiese, 2003; Sherriff, 2010; Smyth, 2011; Sugimoto, 2011; Uçak, 2011; Walcott, 1991; Wu & Chen, 2010). Kushkowsky, Parsons, & Wiese (2003) show that the number of references per thesis varies across disciplines; however, most studies of theses focus on a single discipline and do not compare referencing practices among different types of scholars. One exception is Pancheshnikov (2007) who compared the references of master's theses and faculty publications in biology and found that faculty members use more periodicals and students more monographs. However, it could be argued that these differences could be accounted for the difference in genre (theses vs. monographs) rather than the academic age of the scholars. What are needed are studies that examine patterns across disciplines and by academic strata, using a single genre.

Doctoral mentoring

There have been mixed results on whether or not students receive training on how to conduct a literature search from advisors. Boote and Beile (2005) review literature that suggests doctoral students do not receive this training as advisors assume doctoral students already know how to conduct an adequate search. However, other research suggests that doctoral students receive help from their advisors in this process (e.g., George, Bright, Hurlbert, Linke, St. Clair, & Stein, 2006). Studies also suggest that there is a social aspect to information seeking and use providing evidence that personal, academic and collegial networks are important in seeking information and building reference lists for neophytes (Research Information Network, 2006; Vezzosi, 2008). Furthermore, there is evidence that the disciplinary background of a student's advisor might be related to the selection of interdisciplinary references in the thesis (Sugimoto, Ni, Russell, & Bychowski, 2011). However, few studies have examined large-scale quantitative data to test whether students and faculty vary in their referencing practices.

METHODS

This paper draws on three data sources: the list of all doctoral students enrolled in Quebec's universities between 2000 and 2007, the list of all faculty members in Québec universities over the same period and all 2000-2007 papers with at least one Quebec institutional address in Thomson Reuters' Web of Science database (WoS). The lists of doctoral students and faculty members were matched with authors' names in WoS and disambiguated using both manual and automatic validation. More details on the process can

be found in Larivière (2010). In total, 11,513 papers first-authored by 5,627 doctoral students are compared with 18,988 papers first-authored by 6,372 faculty members.

Each of the approximately 11,000 scientific journals indexed yearly in WoS were categorized into three broad categories for subsequent analysis: health sciences (health), natural sciences and engineering (NSE), and social sciences and humanities (SSH). Publication counts presented in this paper are based on the number of articles to which authors from Québec contributed during the 2000-2007 period.

Two mutually exclusive groups of papers are compared here: Papers first-authored by doctoral students and papers first-authored by faculty members. Only first-authored papers are considered here as, in most disciplines except mathematics and high energy physics where alphabetical order is the norm, the first author is considered as the researcher primarily responsible for the content of the paper (Birnholtz, 2006; Pontille, 2004). Papers are considered as being authored by doctoral students when the first author is enrolled in a doctoral program in one of Quebec's universities during the publication year of the paper or has been enrolled during the year prior to the publication. For faculty members, all 2000-2007 papers are considered.

Several indicators of referencing practices are presented in this paper. The average number of references made is simply an average of all documents cited, including both source and non-source items. The average age of cited literature is obtained by subtracting the publication year of each cited document to the publication year of the citing document, and then averaging these numbers (Larivière, Gingras & Archambault, 2008). In order to reduce the weight of spelling mistakes in publication years, these averages are compiled for a citation window of 100 years. The retrieval of the various types of documents cited is based on the search strategies found in Larivière, Zuccala and Archambault (2008) for theses, Lisée, Larivière and Archambault (2009) for conference proceedings and Larivière et al. (2006) for journal articles. The analysis of self-citations follows the typology of Lawani (1982), which distinguishes between diachronous self-citations (received) and synchronous self-citations (made). While the former is the percentage of self-citations among the citations received by a researcher, the second is the percentage these references account among all material cited by an author. In this paper, the latter method is used. Along these lines, we also use the method of Snyder and Bonzi (1998), which considers as a self-citation any intersection between the authors of the citing authors of a paper and the authors of the cited paper. Given that WoS only indexes the names of co-authors of cited papers that are also source items, this operationalization of self-citation can only be performed for references made to source items (or for citations received by source items), that is, papers indexed in the database. The same concept applies to national citations.

In order to take into account the various citation practices across disciplines and specialties (Moed, 2005), the Impact Factor of each journal — corrected to account for the asymmetry between the numerator and denominator (Archambault & Larivière, 2009) — was divided by the average Impact Factor of journals of the same specialty. In addition, analysis was conducted to evaluate differences in faculty and student referencing patterns by the research level of journal. The research level of journals was determined using the CHI Research classification (Hamilton, 2003), based on the type of research it publishes. It has four levels: i) clinical observation or applied technology; ii) clinical mix or

engineering/technological science; iii) clinical investigation or applied research/targeted basic research; and iv) basic research. Distributions presented in the paper exclude cited journals that do not have a research level assigned. Finally, although the results presented in this paper are based on all doctoral students and faculty in the province of Quebec as well as of all of their WoS-indexed papers—both of which can be considered as population data—we have performed statistical tests of significance (t-tests). All differences were found to be statistically significant at $p < 0.001$, except for the difference in the proportion of these cited for NSE doctoral students and faculty (Figure 2), which was significant at $p < 0.005$, and for the difference between SSH doctoral students and faculty in terms of number of references per paper (left panel of Figure 1), which was not found to be significant. This lack of significance is due to the smaller number of papers published by SSH researchers.

RESULTS

Figure 1 presents the average number of references (left panel) and average age of cited references (right panel), for both doctoral students and faculty members, across the three disciplinary areas. As shown, doctoral students' papers contain, on average, more references than faculty members' papers across all disciplinary areas. The largest difference is in health, where doctoral student papers contain an average of 43.6 references to 38.9 references in the corresponding faculty papers.

The right panel of Figure 1 presents the average age of cited documents for the two groups of papers, using a 100-year citation window. Doctoral student papers cite, on average across all disciplines, documents that are about 10% more recent. For each disciplinary area, the average age of references used by faculty members is about one year older than that of the doctoral student references.

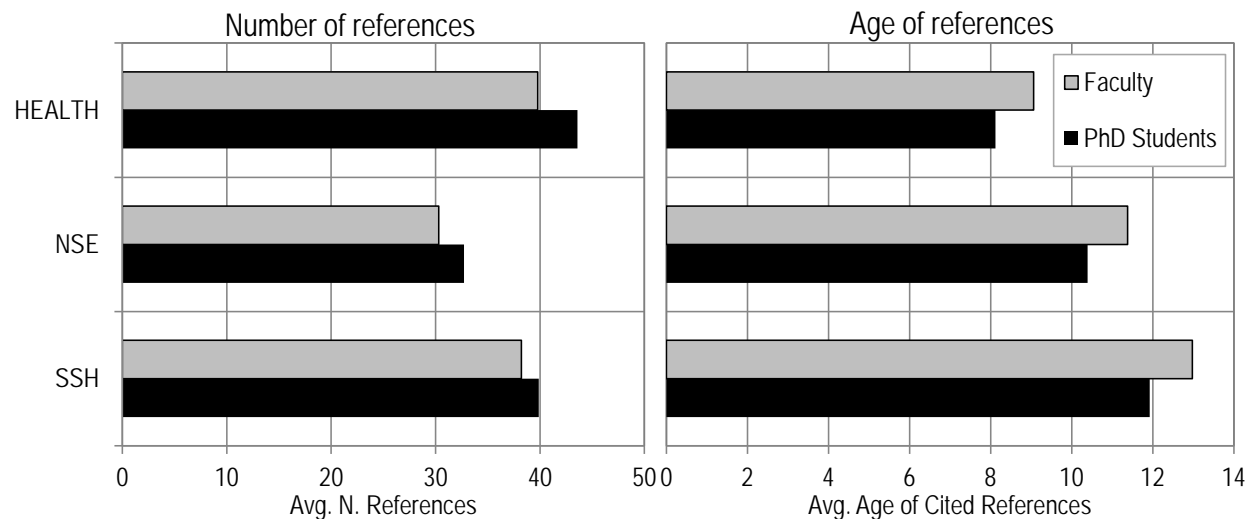


Figure 1. For faculty members and doctoral students: average number of references per paper and average age of cited references.

Figure 2 presents the proportion of cited references that are made to specific types of documents: theses, conference proceedings and journals articles. Since doctoral students cite more documents than faculty members, the following numbers are presented as a percentage of all cited references rather than raw numbers.

Theses are marginally cited sources, by both faculty members and doctoral students; however, they are more likely to be cited by faculty members than doctoral students across all domains. The proportion of conference proceedings cited offers a more nuanced portrait: while doctoral students in health and SSH reference a higher proportion of proceedings than faculty members, the opposite is true in NSE—an area where proceedings account for the largest proportion of all cited material. The percentage of references made to journal articles follow a pattern consistent across all areas, with doctoral students citing a greater proportion of articles than faculty members. This is particularly striking in SSH, where non-serial literature—such as books (Hicks, 1999; Huang & Chang, 2008)—account for a larger proportion of cited literature: serials accounted for 58% of material cited by faculty, yet represented 71% of those cited by doctoral students.

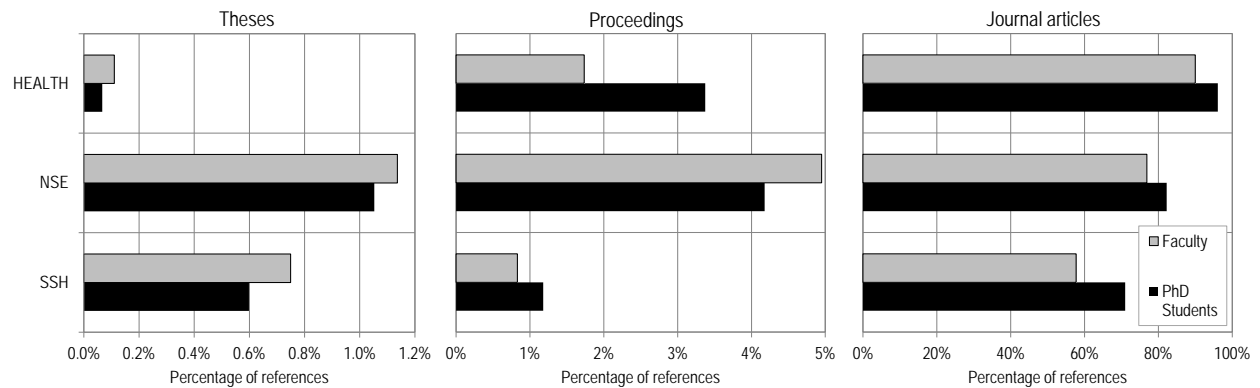


Figure 2. For faculty members and doctoral students: percentage of references made to theses, conference proceedings and journal articles.

Figure 3 presents two aspects of self-citations for both doctoral students and faculty members’ papers: percentage of references made to papers on which any author of the citing papers is one of the authors of the cited paper (self-citations, left panel) and percentage of references on which there is at least one Canadian address (national citations, right panel). This demonstrates that, in each discipline and for both measures (with the exception of national citations in SSH), faculty members are more likely to self-cite or cite national colleagues than doctoral students.

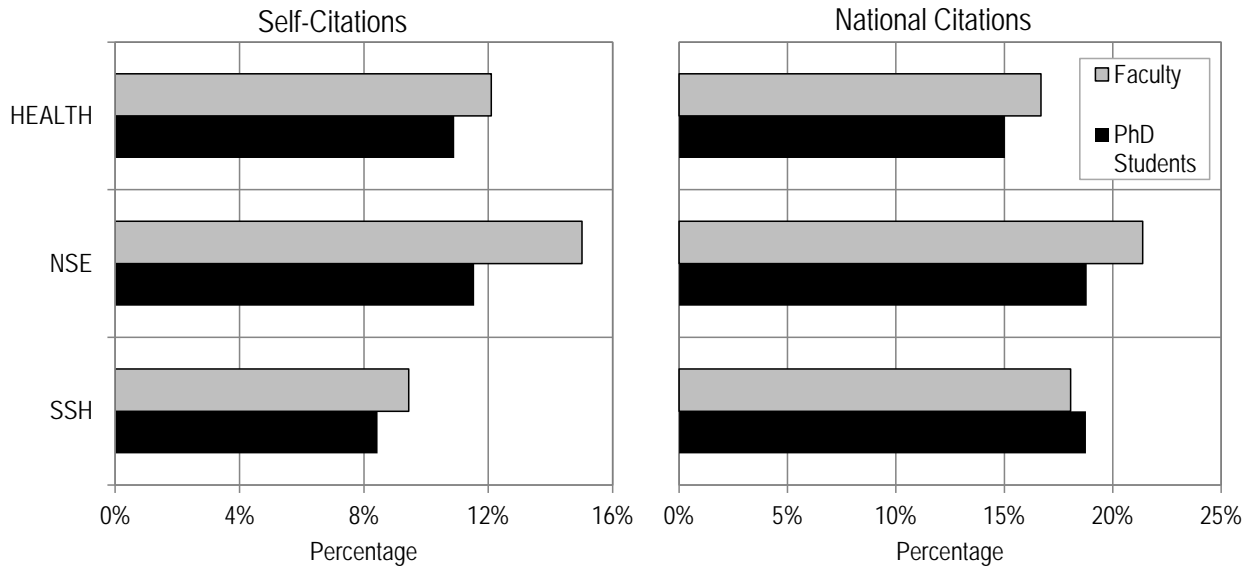


Figure 3. For faculty members and doctoral students: percentage of self-citations (any author) and of national (Canadian) citations.

Figure 4 presents the average impact of cited journals (left panel) and their average research level (right panel). Consistent patterns across disciplinary areas are not exhibited in this analysis. In NSE, there is little difference in the impact of journals referenced by students and faculty. In health, faculty members tend to cite journals with slightly lower impact factors while in SSH, faculty members cite journals with higher impact factors.

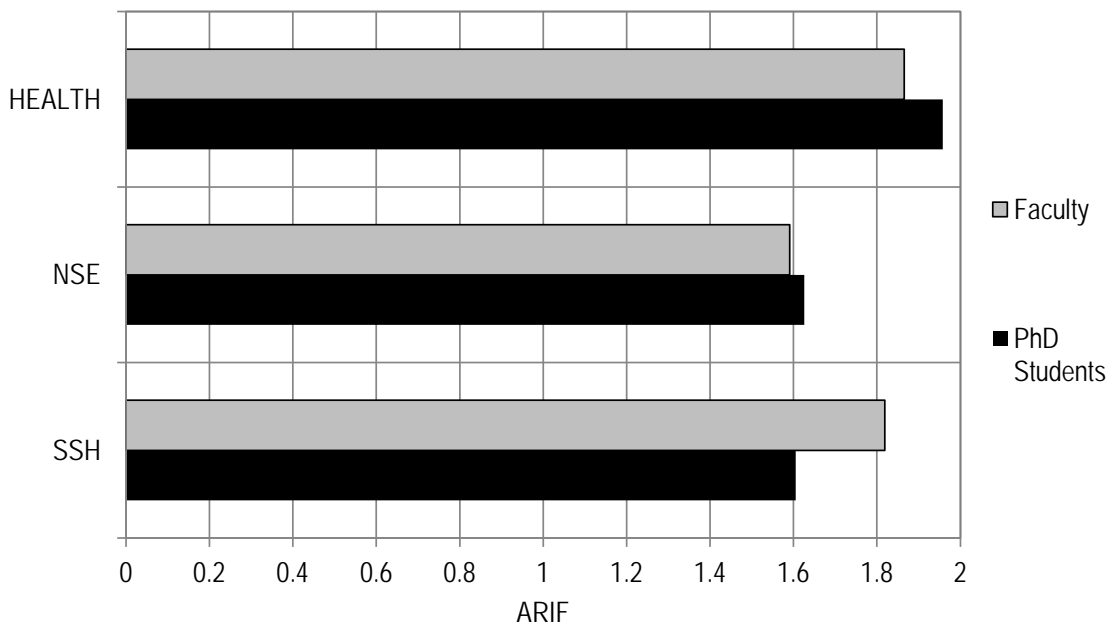


Figure 4. Field-normalized impact of journals (ARIF) cited by faculty members and doctoral students

Finally, Figure 5 provides evidence that, in every discipline, faculty members tend to cite a higher

proportion of clinical/applied research journals (levels I and II) than doctoral students. This difference is quite striking in health research, where 19% and 25% of faculty references are made to level I and II journals respectively, compared to 5% and 1% for doctoral students. A much larger proportion of doctoral students' references are made to journals publishing more basic research (level IV). These differences are quite high in health (57% for doctoral students vs. 34% for faculty members) and in SSH (40% for doctoral students vs. 25% for faculty members). It is worth noting that, while these differences are quite striking in health and SSH, they are much less important in NSE — even if the overall tendency is the same.

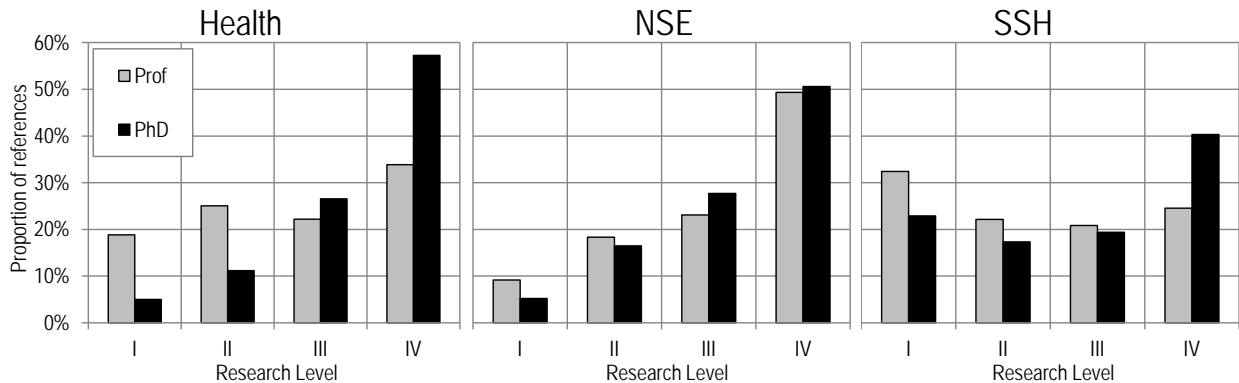


Figure 5. Distribution of the research levels of journals cited by faculty members and doctoral students research level of journals (I = clinical observation or applied technology; II = clinical mix or engineering/technological science; III = clinical investigation or applied research/targeted basic research; and IV = basic research)

DISCUSSION

Significant differences were found between faculty member and doctoral student referencing behavior for the period studied (2000-2007)¹. Discussions of these differences, by research question (RQ), are presented below.

RQ1 and RQ2. The first and second research questions, regarding differences in the number of references made as well as of their age, were answered in the affirmative. Several things can account for the differences observed. First and foremost, while doctoral students are in the process of constructing their identity as researchers, faculty members can be, generally, considered as established researchers. As a consequence, doctoral students are in the process of building their 'own', new set of references, while established researchers are likely to be working in the same research area for a certain amount of time and, hence, re-using the same set of references. Moreover, the fact that doctoral students' papers are likely to be related to their thesis — for which they are necessarily conducting an in-depth screening

¹ T-tests were used to compare the means and proportions and nearly all were statistically significant and $p < 0.001$. The only case where no significant difference was observed is in Panel A of Figure 1 for SSH. This is due to the fact that the N used in this figure is the number of papers (As the metric is an average per paper) instead of the number of references. Given that there are a smaller number of papers in SSH, this insignificance can be explained due to a smaller population than the others.

of the literature of their domain — might explain much of the differences observed, both in terms of the quantity of references and their timeliness. For faculty members, these results are in line with those obtained by Gingras *et al.* (2008) showing that younger researchers' papers cited more recent material than older researchers, with Barnett and Fink (2008) providing a formal model of the relationship between the age of researchers and their referencing patterns.

RQ3. Significant differences were also demonstrated in the types of scholarly genre cited. This study reinforced earlier studies showing that references to theses are low and decreasing (Ghosh, 2009). However, a surprising result was that faculty members tend to cite theses more than doctoral students. One could have expected that doctoral students, themselves being in the process of writing a thesis, would have reviewed carefully other theses on their topic. One potential explanation is that theses are increasingly made of a group of papers, and that students might actually prefer to read and cite the papers instead of the full thesis. The fact that theses cited by faculty members are, on average, more than two years older than those cited by doctoral students seems to support this. Another explanation could be due to the role that citing plays. Theses are considered non-published work and not "prestigious" sources. If citing is, as claimed by Cole (1998) and Wang and White (1999), part of the public demonstration of a doctoral student's mastery of a domain, then dissertations are less likely to be cited (even though they are probably used in the process of understanding and developing the research). Another hypothesis could be related to the research tools used by the students in their bibliographic work, which are less likely to index doctoral theses, especially those that are not online. Lastly, it may be that faculty members are more exposed to theses, given their service on doctoral committees and therefore more likely to refer to these sources. These various hypotheses require further testing.

The proportion of references made to conference proceedings offers a more nuanced portrait: while doctoral students in health and SSH reference a higher proportion of proceedings than faculty members, the opposite is true in NSE. It must be noted that NSE is also the discipline where proceedings account for the largest proportion of all cited material. The fact that doctoral students in health cite about twice as many proceedings might be caused by access, as these documents are more often available online, or by the fact that students may be more likely to attend these conferences. It is interesting to note that in the Research Information Network (2006, p. 86) survey, post-doctoral students rank conference proceedings as the 7th most useful source (with 6.9% of the respondents), while faculty rank them as the 8th most useful (with 11.6% of the respondents) (all sectors combined).

The percentage of references made to journal articles follow a pattern constant across all areas, with doctoral students citing a greater proportion of papers than faculty. This is particularly striking in SSH, where non-serial literature—such as books—account for a larger proportion of cited literature: serials accounted for 58% of material cited by faculty, yet represented 71% of those cited by doctoral students. There is undisputedly an access (convenience) element to this: papers are typically more available than books, and current literature is typically more often found journal articles than in books (Larivière, Gingras & Archambault, 2007).

RQ4. Significant differences were found in terms of the self-citation practices of faculty members and doctoral student. While some authors view self-citations as a problematic aspect of scientific communication (Bras-Amorós, Domingo-Ferrera & Torra 2011; MacRoberts & MacRoberts, 1989), others see it as a normal feature of sustained research activity, where authors draw on their previous research (Glanzél et al., 2006). Similarly, the work of Fowler and Aksnes (2007) shows that self-citations increases the number of citations from others, and that frequent self-citers were not ‘penalized’ by obtaining lower ‘external’ citation rates. Our study demonstrates that, in each discipline and for both measures (with the exception of national citations in SSH), faculty members are more likely to self-cite or cite national colleagues than doctoral students. Of course, having a longer publication history, faculty members are obviously more likely to have a pool of previously published papers they can cite. But given that self-citation measures presented here are drawn on *any* of the authors (both at the cited and citer levels), and that the majority of the papers analyzed in this dataset are co-authored — 97.1% for doctoral students and 80.1% for faculty members — this cannot explain the differences, as other co-authors on the paper could have a long publication history to draw from. On the whole, this suggests that, in their literature reviews, doctoral students read wider and deeper and are more likely than faculty members to cite more distant documents with which the group of authors has no ties.

RQ5. The comparison of these results on impact factors’ of cited papers with the average impact of doctoral student papers found in Larivière (2010) shows that the area where doctoral students’ papers have lower impact (SSH) is also that where they cite journals with a lower impact. This finding seem to agree with the results of Bornmann, de Moya Anegón and Leydesdorff (2010), which tested the Ortega hypothesis² and showed that high impact papers were more likely to cite other high-impact papers. On the other hand, this does not hold true for NSE and health, as doctoral students’ papers in NSE have, relatively, more impact in NSE than health, even though students in health have, on average, cited more high impact journals. Taken globally, this comparison of the impact of cited journals also suggest that doctoral students in the sciences (NSE and health) have, on average, a better knowledge of the important sources of their field than doctoral students in SSH. This might be explained by the fact that the sciences typically have lower task-uncertainty (Whitley, 1984) and one dominant paradigm (Kuhn, 1962) compared with SSH, which, in turn, might translate in a better knowledge of the quality of the sources.

RQ6. The final finding suggests that the norm is still to cite basic research instead of applied for doctoral work. Hence, despite the rise in percentage of doctoral students who perform their doctoral research in a non-university—and thus more applied—setting (Gemme, 2009; Gemme & Gingras, 2008; Thune, 2009), it seems that the research performed by doctoral students is still quite basic.

CONCLUSION AND FUTURE DIRECTIONS

This paper has compared several components of doctoral students’ and faculty members’ referencing behavior, based on the entire population of WoS-indexed research papers first-authored by either group

² The Ortega hypothesis — named after the Spanish philosopher José Ortega y Gasset — suggests that high impact research leans on average research.

between 2000 and 2007. Significant differences in referencing behavior were identified between these groups. Returning to the notion of communities of practice, we might account for these differences by status: doctoral students are neophytes in an early stage of initiation who will slowly acculturate to the referencing practices of faculty members. However, this assumes a degree of stasis in referencing patterns, assuming that young scholars across all time periods behave in the same ways relative to their older colleagues. Given that the majority of studies have been focused on a small time period, more research is needed to see how variations in time, discipline, technology, and other sociotechnical factors might change the relative differences between these groups of scholars.

This study reinforced earlier findings that younger scholars tend to cite more recent literature. It is worth noting that Nobel laureates tend to cite more like their younger colleagues (Zuckerman & Merton, 1973). This implies a potential connection between innovation and awareness of recent literature. A more recent study of Spanish scientists noted that the younger scholars in each area were those with the highest impact (Costasa, van Leewen, & Bordons, 2010). However, Zuckerman and Merton's (1973, p. 515) comment remains as true today as it was in 1973: "evidence on age differentials in receptivity to new ideas in science remains thin and uncertain." Additional large-scale quantitative analyses are still needed to more fully understand the relationship between age, information seeking, and referencing behavior.

References

- Archambault, É., Larivière, V. (2009) History of the journal impact factor: contingencies and consequences. *Scientometrics*, 79(3), 639-653.
- Barett, A. (2005). The information-seeking habits of graduate student researchers in the humanities. *The Journal of Academic Librarianship*, 31(4), 324-331.
- Barnett, G.A., & Fink, E.L. (2008). Impact of the internet and scholar age distribution on academic citation age. *Journal of the American Society for Information Science and Technology*, 59(4), 526-534.
- Birnholtz, J. (2006). What does it mean to be an author? The intersection of credit, contribution and collaboration in science. *Journal of the American Society for Information Science and Technology*, 57(13), 1758-1770.
- Boote, D.N., & Beile, P. (2005). Scholars before researchers: On the centrality of the dissertation literature review in research preparation. *Educational Researcher*, 34(6), 3-15.
- Bornmann, L., & Daniel, H.-D. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, 64(1), 45-80.
- Bornmann, L., de Moya Anegón, F., Leydesdorff, L. (2010). Do scientific advancements lean on the shoulders of giants? A bibliometric investigation of the Ortega Hypothesis, *PLoS ONE*, 5(10), e13327.
- Bras-Amorós, M., Domingo-Ferrer, J., Torra, V. (2011). A bibliometric index based on the collaboration distance between cited and citing authors. *Journal of Informetrics*, 5(2), 248-264.
- Brooks, T.A. (1986). Evidence of complex citer motivations. *Journal of the American Society for Information Science*, 37(1), 34-36.
- Budd, J.M. (1990). Higher education literature: Characteristics of citation patterns. *The Journal of Higher Education*, 61(1), 84-97.
- Buttler, L. (1999). Information sources in library and information science doctoral research. *Library & Information Science Research*, 21(2), 227-245.

- Case, D.O. (2007). *Looking for information: a survey of research on information seeking, needs, and behavior*. 2nd Edition. London: Academic Press.
- 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 and Technology*, 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.
- Cole, C. (1997). Information as process: The difference between corroborating evidence and "information" in humanistic research domains. *Information Processing & Management*, 33(1), 55-67.
- Cole, C. (1998). Information acquisition in history Ph.D. students: Inferencing and the formation of knowledge structures. *Library Quarterly*, 68(1), 33-54.
- Cole, C. (2000). Name collection by Ph.D. history students: Inducing expertise. *Journal of the American Society for Information Science*, 51(5), 444-455.
- Conkling, T.W., Harwell, K.R., McCallips, C., Nyana, S.A., & Osif, B.A. (2010). Research material selection in the pre-web and post-web environments: An interdisciplinary study of bibliographic citations in doctoral dissertations. *The Journal of Academic Librarianship*, 36(1), 20-31.
- Costas, R., van Leeuwen, T.N., & Bordons, M. (2010). A bibliometric classificatory approach for the study and assessment of research performance at the individual level: The effects of age on productivity and impact. *Journal of the American Society for Information Science & Technology*, 61(8), 1564-1581.
- Creaser, C., Oppenheim, C., & Summers, M.A.C. (2011). What do UK academics cite? An analysis of references cited in UK scholarly outputs. *Scientometrics*, 86, 613-627.
- Cronin, B. (1981). Agreement and divergence on referencing practice. *Journal of Information Science*, 3, 27-33.
- Davenport, E., & Hall, H. (2002). Organizational knowledge and communities of practice. In B. Cronin (Ed.), *Annual Review of Information Science and Technology*, 36 (pp. 171-227). Medford, NJ: Information Today.
- Davies, E. (2005). Communities of practice. In K.E. Fisher, S. Erdelex, & L. McKechnie, (Eds.), *Theories of information behavior* (pp.103-106). Medford, N.J.: Information Today.
- de Solla Price, D.J., & Beaver, D. (1966). Collaboration in an invisible college. *American Psychologist*, 21(11), 1011-1018.
- Delgadillo, R., & Lynch, B.P. (1999). Future historians: Their quest for information. *College & Research Libraries*, 60(3), 245-259.
- Feyereisen, P., & Spoiden, A. (2009). Can local citation analysis of master's and doctoral theses help decision-making about the management of the collection of periodicals? A case study in psychology and education sciences. *The Journal of Academic Librarianship*, 35(6), 514-522.
- Fleming-May, R., & Yuro, L. (2009). From student to scholar: The academic library and social science PhD students' transformation. *portal: Libraries and the Academy*, 9(2), 199-221.
- Fowler, J.H. & Aksnes, D.W. (2007). Does self-citation pay? *Scientometrics*, 72, 427-437.
- Frandsen, T.F., & Nicolaisen, J. (2012). Effects of academic experience and prestige on researchers' citing behavior. *Journal of the American Society for Information Science and Technology*, 63(1), 64-71.
- Gao, S.-J., Guo, J.X., & Duan, X.-L. (2011). The study of library use and document gathering behavior: A survey of geomatics faculty at Wuhan University, China. *Library Collections, Acquisitions, & Technical Services*, 35, 19-28.
- Ge, X. (2010). Information-seeking behavior in the digital age: A multidisciplinary study of academic researchers. *College & Research Libraries*, 71(5), 435-455.

- Gemme, B. (2009). *The outside within: Heteronomy in the training of forest researchers*. Doctoral thesis: University of British Columbia.
- Gemme, B., and Gingras Y. (2008). The new production of researchers, In Chan, A., and Fischer, D. (Eds.) *The exchange university*. Vancouver: UBC Press, pp. 70–89.
- George, C., Bright, A., Hurlbert, T., Linke, E.C., St. Clair, G., & Stein, J. (2006). Scholarly use of information: Graduate students' information seeking behavior. *Information Research-An International Electronic Journal*, 11(4), 272.
- Ghosh, M. (2009). E-theses and Indian academia: A case study of nine ETD digital libraries and formulation of policies for a national service. *The International Information & Library Review*, 41, 21-33.
- Gingras, Y., Larivière, V., Macaluso, B., & Robitaille, J.P. (2008). The effects of aging on researchers' publication and citation patterns. *PLoS ONE*, 3(12), e4048. arXiv:0810.4292.
- Girves, J.E., & Wemmerus, V. (1988). Developing models of graduate student degree progress. *The Journal of Higher Education*, 59(2), 163-189.
- Glanzel, W., Debackere, K., Thijs, B. & Schubert, A. (2006) A concise review on the role of author self-citations in information science, bibliometrics and science policy. *Scientometrics*, 67, 263-277.
- Gooden, A.M. (2001). Citation analysis of chemistry doctoral dissertations: An Ohio State University case study. *Issues in Science and Technology Librarianship*. Retrieved online at: <http://www.istl.org/01-fall/refereed.html>
- Hamilton, K. (2003) *Subfield and Level Classification of Journals* (CHI Report No. 2012-R). Cherry Hill, NJ: CHI Research.
- Haring, M.J. (1999). The case for a conceptual base for minority mentoring programs. *Peabody Journal of Education*, 74(2), 5-14.
- Hellqvist, B. (2010). Referencing in the humanities and its implications for citation analysis. *Journal of the American Society for Information Science and Technology*, 61(2), 310-318.
- Hicks, D.M. (1999). The difficulty of achieving full coverage of international social science literature and the bibliometric consequences. *Scientometrics*, 44(2), 193-215.
- Hjorland, B., & Albrechtsen, H. (1995). Toward a new horizon in information science: Domain-analysis. *Journal of the American Society for Information Science*, 46(6), 400-425.
- Huang, M.H & Chang, Y.W. (2008). Characteristics of research output in social sciences and humanities: From a research evaluation perspective. *Journal of the American Society for Information Science and Technology*, 59(11), 1819–1828.
- Jamali, H.R. & Nicholas, D. (2006). Communication and information-seeking behavior of PhD students in physicists and astronomy. *Proceedings of the American Society for Information Science and Technology*, 43(1).
- Keat, Y.C., & Kiran, K. (2008). Citation study of library and information science dissertations for collection development. *Malaysian Journal of Library & Information Science*, 13(2), 29-47.
- Khosrowjerdi, M., & Iranshahi, M. (2011). Prior knowledge and information-seeking behavior of PhD and MA students. *Library & Information Science Research*, 33, 331-335.
- Kidd, J.S. (1990). Measuring referencing practices. *Journal of the American Society for Information Science*, 41(3), 157-163.
- Korobili, S., Malliari, A., & Zapounidou, S. (2011). Factors that influence information-seeking behavior: The case of Greek graduate students. *The Journal of Academic Librarianship*, 37(2), 155-165.
- Kuhn, T.S. (1962). *The structure of scientific revolution*. Chicago: Chicago University Press.
- Kumar, H.A., & Dora, M. (2011). Citation analysis of doctoral dissertations at IIMA: A review of the local use of journals. *Library Collections, Acquisitions, & Technical Services*, 35, 32-39.
- Kushkowsky, J.D. (2005). Web citation by graduate students: A comparison of print and electronic thesis. *portal: Libraries and the Academy*, 5(2) : 259–276.

- Kushkowski, J.D., Parsons, K.A. & Wiese, W.H. (2003). Master's and doctoral thesis citations: Analysis and trends of a longitudinal study. *portal: Libraries and the Academy*, 3(3), 459–479.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.
- Larivière, V. (2010). *A bibliometric analysis of Quebec's PhD students' contribution to the advancement of knowledge*. Ph.D. thesis, McGill University.
- Larivière, V., Archambault, É., Gingras, Y. (2007) Long-term patterns in the aging of the scientific literature, 1900–2004, in Torres-Salinas, D. and Moed, H.F. (eds) *Proceedings of the 11th International Conference of the International Society for Scientometrics and Informetrics (ISSI)*, Madrid: CSIC, pp. 449-456.
- Larivière, V., Archambault, É., Gingras, Y. (2008) Long-term variations in the aging of scientific literature: from exponential growth to steady-state science (1900-2004). *Journal of the American Society for Information Science and Technology*, 59(2), 288-296.
- Larivière, V., Archambault, É., Gingras, Y., & Vignola–Gagné, É. (2006). The place of serials in referencing practices: Comparing natural sciences and engineering with social sciences and humanities. *Journal of the American Society for Information Science and Technology*, 57(8), 997–1004.
- Larivière, V., Macaluso, B., Archambault, É., & Gingras, Y. (2010). Which scientific elites? On the concentration of research funds, publications and citations. *Research Evaluation*, 19(1), 45-53.
- Larivière, V., Zuccala, A., Archambault, É. (2008) The declining scientific impact of thesis: Implications for electronic thesis and dissertation repositories and graduate studies. *Scientometrics*, 74(1), 109-121.
- Lawani, S.M. (1982). On the heterogeneity and classification of author self-citations. *Journal of the American Society for Information Science*, 33(5), 281–284.
- Liao, Y., Finn, M., & Lu, Jun, L. (2007). Information-seeking behavior of international graduate students vs. American graduate students: A user study at Virginia tech 2005. *College & Research Libraries*, 68(1), 5-25.
- Line, M.B. (1971). The information uses and needs of social scientists: An overview of INFROSS. *Aslib Proceedings*, 23(8), 412-434.
- Lisée, C., Larivière, V., Archambault, É. (2008) Conference proceedings as a source of scientific information: A bibliometric analysis. *Journal of the American Society for Information Science and Technology*, 59(11), 1776–1784.
- Luker, K. (2008). *Salsa dancing into the social sciences: research in an age of info-glut*. Cambridge: Harvard University Press.
- 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.
- Meho, L.I., & Haas, S.W. (2001). Information-seeking behavior and use of social science faculty studying stateless nations: A case study. *Library & Information Science Research*, 23, 5-25.
- Meho, L.I., & Tibbo, H.R. (2003). Modeling the information-seeking behavior of social scientists: Ellis's study revisited. *Journal of the American Society for Information Science and Technology*, 54(6), 570-587.
- Milojević, S. (2011). Citation practices in the field of library and information science. *Annual Meeting of the American Society for Information Science and Technology (ASIST 2011) Workshop: Metrics 2011 - Symposium on Informetric and Scientometric Research (Presentation) - New Orleans, LA*.
- Milojević, S., Sugimoto, C.R., Yan, E., & Ding, Y. (2011). The cognitive structure of library and information science: Analysis of article title words. *Journal of the American Society for Information Science and Technology*, 62(10), 1933-1953.
- Moed, H.F. (2005). *Citation analysis in research evaluation*. Dordrecht: Springer.

- Nederhof, A.J. (2011). A bibliometric study of productivity and impact of modern language and literature research. *Research Evaluation*, 20(2), 117-129.
- 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.
- Pancheshnikov, Y. (2007). A comparison of literature citations in faculty publications and student theses as indicators of collection use and a background for collection management at a university library. *The Journal of Academic Librarianship*, 33(6), 674-683.
- Pontille, D. (2004). *La signature scientifique: Une sociologie pragmatique de l'attribution*. Paris: CNRS Éditions.
- Research Information Network. (2006). *Discovery services: user behaviour, perception and needs. A study commissioned by the Research Information Network*. Retrieved online at: <http://www.rin.ac.uk/our-work/using-and-accessing-information-resources/researchers-and-discovery-services-behaviour-perc>
- Sherriff, G. (2010). Information use in history research: A citation analysis of master's level theses. *portal: Libraries and the Academy*, 10(2), 165-183.
- Silverman R.J. (1985). Higher education as a maturing field? Evidence from referencing practices. *Research in Higher Education*, 23(2), 150-183.
- Smyth, J.B. (2011). Tracking trends: Students' information use in the social sciences and humanities, 1995-2008. *portal: Libraries and the Academy*, 11(1), 551-573.
- Snyder, H. and Bonzi, S. (1998). Patterns of self-citation across disciplines (1980-1989). *Journal of Information Science*, 24, 431-435.
- Sugimoto, C.R. (2011). Looking across communicative genres: A call for inclusive indicators of interdisciplinarity. *Scientometrics*, 86(2), 449-461. doi: 10.1007/s11192-010-0275-8
- Sugimoto, C.R., Ni, C., Russell, T.G., & Bychowski, B. (2011). Academic genealogy as an indicator of interdisciplinarity: An examination of dissertation networks in Library and Information Science. *Journal of the American Society for Information Science & Technology*, 62(9), 1808-1828.
- Sugimoto, C.R. (2012). Are you my mentor? Identifying mentors and their roles in LIS doctoral education. *Journal of Education for Library & Information Science*, 53(1), 2-19.
- Talja, S., & Maula, H. (2003). Reasons for the use and non-use of electronic journals and databases: A domain analytic study in four scholarly disciplines. *Journal of Documentation*, 59(6), 673-691.
- Uçak, N.O. (2011). Information use in art: A citation analysis of sources used in art theses in Turkey. *The Journal of Academic Librarianship*, 37(1), 77-81.
- Vezzosi, M. (2009). Doctoral students' information behavior: An exploratory study at the University of Parma (Italy). *New Library World*, 111(1/2), 65-80.
- Walcott, R. (1991). Characteristics of citations in geoscience doctoral dissertations accepted at United States academic institutions 1981-1985. *Science & Technology Libraries*, 12(2), 5-16.
- Wang, P., & White, M. D. (1999). A cognitive model of document use during a research project. Study II. Decisions at the reading and citing stages. *Journal of the American Society for Information Science and Technology*, 50(2), 98-114.
- White, H.D. (2001). Authors as citers over time. *Journal of the American Society for Information Science & Technology*, 52(2), 87-108.
- Whitley, R. (1984). *The Intellectual and Social Organization of the Sciences*. Oxford: Oxford University Press.
- Wu, M.-D., & Chen, S.-C. (2010). The impact of electronic resources on humanities graduate student theses. *Online Information Review*, 34(3), 457-472.

Zuckerman H., & Merton R.K. (1973). Age, aging and age structure in science. In: Storer, N.W. (Ed.) *The Sociology of Science: Theoretical and empirical investigations*. Chicago: Chicago University Press. pp. 497–559.