informationresearch

PUBLISHED QUARTERLY BY THE UNIVERSITY OF BORÅS, SWEDEN

VOL. 23 NO. 2, JUNE, 2018

Use of Web search engines and personalisation in information searching for educational purposes

Sara Salehi, Jia Tina Du, and Helen Ashman.

Introduction: *Students increasingly depend on Web search for educational purposes. This causes concerns among education providers as some evidence indicates that in higher education, the disadvantages of Web search and personalised information are not justified by the benefits.*
Method: *One hundred and twenty university students were surveyed about their information-seeking behaviour for educational purposes. We also examined students' information access while using Web search, through twenty-eight one-on-one study sessions.*
Analysis: *Survey participants ranked their preference towards different information resources on a 5-point Likert scale. Given equal exposure to the first five standard pages of the search results during the study sessions, students' explicit and implicit feedback was used to evaluate the relevance of the search results.*
Results: *First, most participating students declared that they use Google search engine as their primary or only information-seeking tool. Second, about 60% of the clicked result links during the study sessions were located in pages 2+ of the search results without personalisation influencing the relevance of the top-ranked search results. In real-life scenarios pages 2+ of the search results receive only ~10% of the clicks. Students also expressed more satisfaction with the relevance of non-personalised over personalised search results. These differences presented a missed information opportunity, an opportunity bias, for students.*

Introduction

Google is the world's most popular search engine, with 90% of the worldwide search engine market share from 2010 to 2016 (Statista, 2016). In 2009 Google started providing personalised search results for all users, even those without a Google account (Horling and Kulick, 2009). Personalised search apparently benefits users in ways such as better understanding the real meaning of ambiguous queries, retrieving locally-relevant results and ultimately an enhanced search experience (Hannak *et al.*, 2013). However, Web search engines are used so often and routinely that we normally do not question their long-term implications.

Some of the main concerns surrounding personalised information, particularly in academic settings, are loss of serendipity, capability and deep learning (Ashman, Brailsford and Brusilovsky, 2009; Ashman *et al.*, 2014; Pariser, 2011). Google gives users the sense that the Internet is a part of their own cognitive tool set and makes them feel more capable and knowledgeable than before, when in fact users' reliance on Web search engines makes them far less educated about the world around them (Sparrow, Liu and Wegner, 2011). This is especially concerning when it comes to research and higher education. Despite advances in eLearning systems, students still depend heavily on general-purpose search engines to find academic information (McClure and Clink, 2009; Purdy, 2012) and Google their way through their research projects (Brabazon, 2016)

Additionally, users' Web search behaviour is highly, if unwittingly, prejudiced. Users have a strong faith in what the search engines think is best and that affects their choice of search results; furthermore, they tend to choose positive information over negative and prefer higher-ranked search results even if they are less relevant to their need. These biases ultimately change the search results based on what a user prefers, as search engines tend to please the searchers by providing them with familiar concepts in their domain of interest while leaving out potentially important but less-agreeable information. This is known as the filter bubble effect (Pan *et al.*, 2007; Pariser, 2011; White, 2013).

The study reported here will help Web search engine researchers and designers to bridge the semantic gap between the user and the system by highlighting the positive and mediating the risks of frequent use of personalised Web search in higher education. Personalised search refers to Web search experiences that are tailored to an individual by integrating information about the individual beyond the provided search query. By contrast, non-personalised search

refers to the search experiences that were comparatively independent of individual's personal characteristics. The objectives of this study and the corresponding experiments consist of two parts that address the following research questions:

1. How do Web search engines compare to other academic information resources in terms of students' patterns of use and the importance of information source?
2. How do Web search biases influence students' interactions with search engines? This is in terms of exposure to information and opportunity to access what students find to be more interesting or relevant to their search query.

Part I addresses the first research question by learning about the role and importance of Web search engines in higher education compared to conventional academic information resources such as libraries. We asked 120 higher education students from different groups of undergraduate, research and coursework postgraduate students from a variety of subjects to complete a survey about their academic information-seeking behaviour and preferences. The aim was to learn what proportion of university students use search engines as their primary or only source of information, how important they find search engines in comparison with other academic information sources and which search engine they predominantly use. Part I explains the design of the survey, its scope, participants, and the results.

Part II addresses the second research question by investigating how Web search biases influence students' interactions with search engines and whether search engine ranking strategies affected students' information searching process. This is in terms of exposure to information and opportunity to access the information students find most relevant to their search query. During semi-controlled, one-on-one study sessions, twenty-eight higher education students completed five Web search tasks requiring them to find information on a particular topic, while they were provided with fifty search result links, instead of the default ten, in the first page of the search results. They were asked to read or at least skim through all fifty titles and abstracts before choosing the links they found most relevant. The data come from recording their screen activity, researcher's observation notes and a questionnaire completed after each search task.

The research in this paper extends the work presented in (Salehi, Du and Ashman, 2015) by conducting further analysis on the survey data in Part I and performing further experiments on students' information opportunity in Part II. Information opportunity refers to higher education students' exposure and access to the information they need and find relevant. It describes their chance to see and click on the information that is most relevant and useful to their search query.

Part I: Survey on students' information-seeking behaviour

A survey was designed to observe the information-seeking behaviour of university students and the role that general-purpose search engines play. The survey comprised nine questions including demographic information such as age, sex, field of study and level of education. Then students indicated which search engine they predominantly use and were asked to consider a set of options as information resources for the purpose of completing their academic tasks such as assignments, reports or research. They ranked each option based on:

- How important they find the information resource.
- How often they use each of the information resources.
- How often they use each of the information resources as their first point of enquiry. For instance, if they use Google to search for a subject and choose a link from the search results, the first point of enquiry is Google, not the Website they were directed to by the search results.

Ranking was based on a Likert scale of 1 (lowest) to 5 (highest). General-purpose search engines were named among other options, as we did not want the participants to know the exact purpose of the survey in case it led to biased answers. These other options included library resources, online database and journal subscriptions, textbooks, lecture notes and discipline-specific academic online sources.

A total of 120 students from the University of South Australia were recruited to take part in the survey. Participants included undergraduate (58%), research postgraduate and coursework postgraduate (42%) students from a variety of subjects and attendance statuses (Figure 1). Involving students from a cross-section of the university makes it possible to find if there are discipline-specific differences in search engine usage.

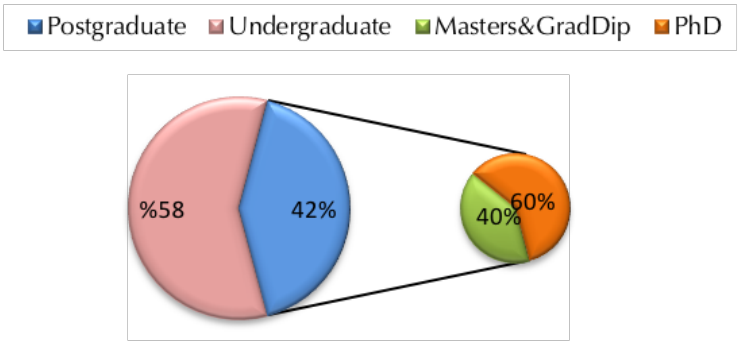

Figure 1: Participant distribution

Survey results

Almost all participants mentioned Google as the search engine they primarily use (Figure 2). This is consistent with previous studies on the information-seeking behaviour of students (Du and Evans, 2011; McClure and Clink, 2009; Purdy, 2012).

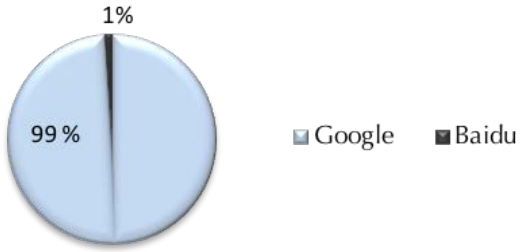

Figure 2: Participants' search engine usage

Next, we asked students to rank the following information resources in terms of importance, frequency of use in general and as the first point of academic information enquiries:

- Library resources (including online databases and journal subscriptions)
- Textbooks
- General-purpose search engines
- Lecture notes
- Discipline-specific and academic online sources

Students chose their answers on a 5-point Likert scale from 1 (lowest) to 5 (highest). The average point for each question (which lies between 1 and 5), reported in Figures 3 and 4, was calculated as follows, where n is the number of participants choosing a specific point and N is the number of participants answering the question:

$$\text{Average point} = (n\times5) + (n\times4) + (n\times3) + (n\times2) + (n\times1) / N$$

Figures 3 and 4 show the average Likert points for importance, frequency of use in general and as the first point of academic enquiries for each academic source and general-purpose search engines. The results indicate to what degree university students depend on general-purpose search engines for their education. Since the differences in average points for other groupings of students were not significant, the figures only show the average points for undergraduate and postgraduate, both research and coursework, students.

Figure 3 shows that not only do higher education students prefer Google to other search engines, they also depend on it as a reliable source of information for conducting assignments and research. More than 83% of participants found search engines an important or very important source of academic information. Figure 3 shows undergraduate students rely on general-purpose search engines, specifically Google, as their most important and reliable academic information resource. Understandably, they find lecture notes and textbooks just about as essential as Google, with all these three options having average Likert points greater than 4. On the other hand, postgraduate students find library resources, including scientific journals and databases, more reliable than search engines, with respective average Likert points of 4.5 and 4.

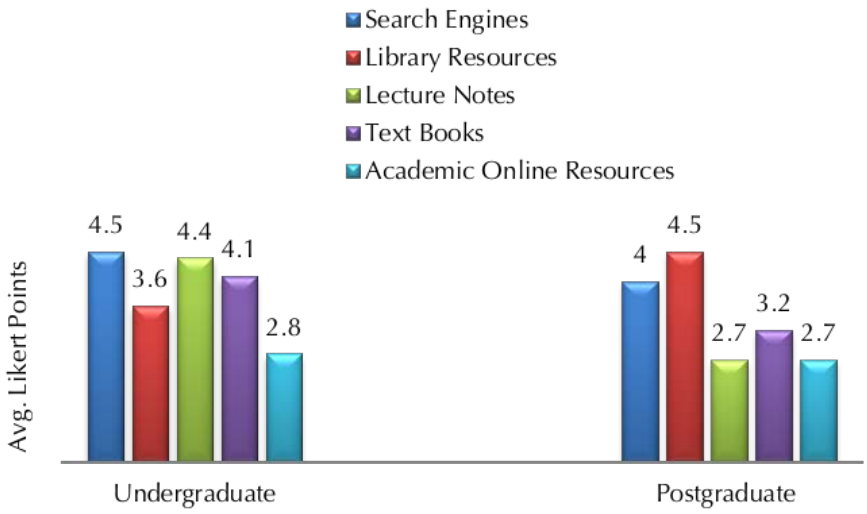

Figure 3: Importance of academic information resources

However, postgraduates' behaviour is not consistent with what they believe. They are encouraged to look beyond search engines and to depend more on academic journals and library resources for their research, but in reality they use the Google search engine more than the designated academic resources with the average Likert point of 4.6. This is demonstrated in Figure 4, where we asked the participants to indicate how often they use each of the information resources in general and also as their starting point of enquiry for seeking information on the topic of their assignments or research. Their responses were ranked based on a 5-point Likert scale: 1 (never) to 5 (daily).

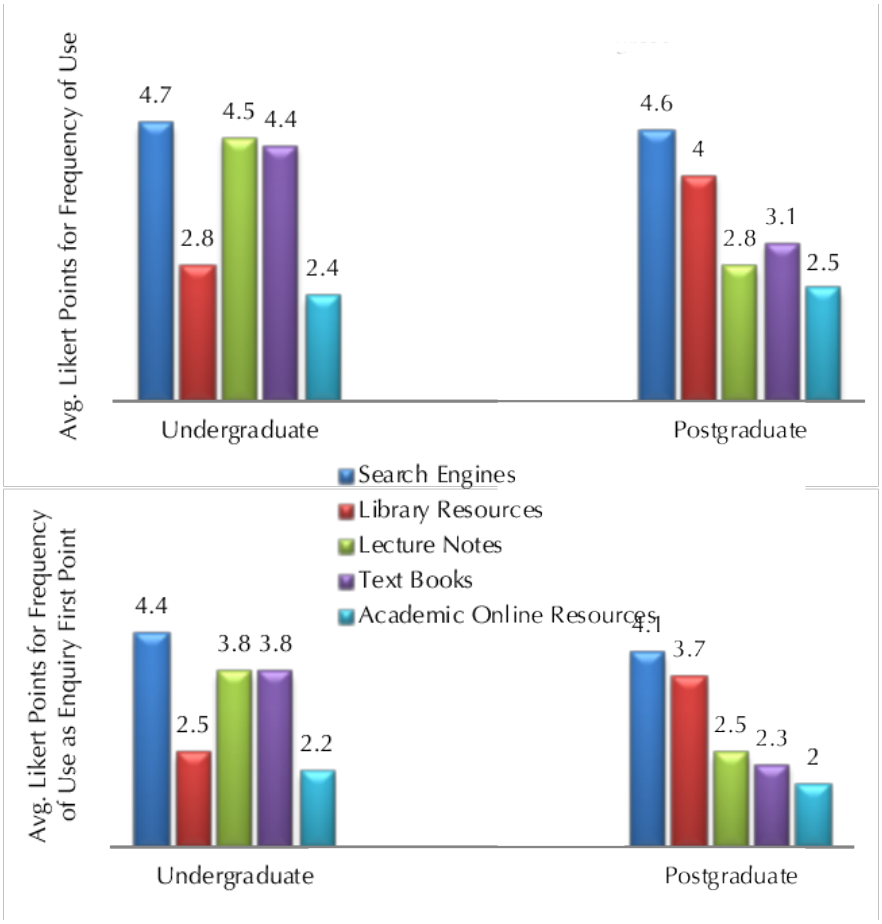

Figure 4: Frequency of use of academic information resources

The average point for how frequently all participants use search engines for educational purposes was 4.7, indicating that most of the 120 university student participants use general-purpose search engines daily as a learning tool. Although we expected postgraduate students to depend less on search engines as a learning resource compared to undergraduate students,

the survey results did not show a meaningful dissimilarity between them with the overall average point being approximately 4.5. This suggests that while postgraduate students are aware of and able to use more reliable educational resources provided by the university, information-seeking habits are entrenched by their undergraduate years and are not changed as their education progresses. The results also indicated an average point of 4.2 for the frequency in which participants use general-purpose search engines as their starting point of enquiry for seeking academic information. Therefore, at least once a week our participants start their self-directed learning sessions with a search engine. Once more, we did not identify any significant differences between different groups of participants.

In summary, for over 80% of higher education students, the general-purpose Google search engine is the most important, relied-on and frequently-used source of academic information. Lecture notes and text books, both with a 3.8 average Likert point, and library resources with the average Likert point of 3.7 come second in terms of regularity of use for undergraduate and postgraduate students, respectively. Discipline-specific and academic resources, e.g., the ACM digital library, Google Scholar and Medline, were less popular even amongst postgraduate research students with average Likert points of less than 3 in importance and frequency of use.

The results of this survey were, to some extent, predictable. Although we anticipated that, in general, students would rely on the general-purpose Google search engine for educational purposes, we did not expect them to so often overlook free, well-known and recommended educational resources such as scientific journals, databases or even Google Scholar in favour of general-purpose search engines, especially given that these search engines are not tailored to the provision of academic-quality research results. We also did not anticipate the degree of this dependency to be relatively the same for all groups of students. It was expected that postgraduate students, especially at doctoral level, would not consider Google search engine as one of their primary research resources, particularly because they are advised how to use specialist research resources at this university.

The effect of students' reliance on Google on their learning process is important, since both traditional and distance education are increasingly self-driven. There is the possibility of this phenomenon creating new opportunities to improve learning or, conversely, causing unexpected pedagogical problems, such as providing different search results based on academically-inappropriate personalisation factors such as location. This motivated the next experiment. Part I showed that students use and trust Web search engines to a high degree. Part II investigates whether this dependency on Web search engines facilitates or hinders students' opportunity to access relevant and high quality information.

Part II: Web search and students' information opportunity

Considering higher education students' significant use of Web search and high visibility of the first page of the search results receiving approximately 90% of the overall search result clicks (Advanced Web Ranking, 2017), the aim of the experiment was to:

- Determine whether personalisation in search helps students by understanding their information needs and placing that information high within the first page where it is more likely to be seen and clicked.
- Verify whether the first page indeed contains the most relevant search results based on students' information needs. This could be the reason behind it receiving almost all the clicks. The alternative is that the most relevant search results are not placed within the first page. In this case, as most Web search users do not look beyond the first page of the search results, users' own position bias and search engines' way of ranking the search results would compromise their information opportunity in real-life scenarios.

Experiment design and implementation

Twenty-eight higher education students from the University of South Australia participated in an experiment conducted over four weeks in late 2015. They included undergraduate (46%) and postgraduate (54%) students from the Schools of Education and Information Technology. All students were required to have a Google account that they had been using regularly for at least a year before attending the study session. The study sessions had no time restrictions and ran, on average, for between forty-five and sixty minutes.

Each student was to complete five Web search tasks to find information on a given topic. Although the tasks were pre-defined, each task would end when the students felt that they had either collected the needed information or searched enough on the topic. They were free to submit as many search queries and choose as many search results as they wished. The

researcher took observational notes and recorded students' screen activity during the sessions for capturing real-time interactions.

Having one group of students complete the tasks using personalised search and the other using non-personalised search would increase the noise in the data (e.g., the difference in the outcome could be an indicator of the difference between the participants' attitudes and skills and not necessarily the difference between search results). This way, the difference in the collected data from personalised and non-personalised experiments would depend on student's individual characteristics and search skills as well as the relevance of the search results. Therefore, as shown in Figure 5, all the participating students in this experiment were unknowingly switched between Google personalised and non-personalised search to complete the tasks. This method ensured that each task was completed using both personalised and non-personalised search. Also, each student experienced and evaluated both personalised and non-personalised search results.

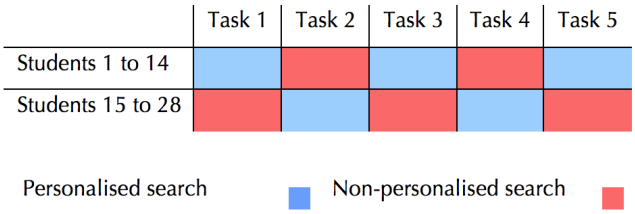

Figure 5: Task distribution

Task design and instructions were identical for all tasks, requiring the participant to state a fact, which he or she finds interesting about the following topics:

- Deepwater Horizon
- Cholesterol
- Immersive environments
- Artificial intelligence
- Pedagogical issues

Before starting the search tasks, the search settings were altered so students would receive fifty, instead of the standard ten, result links in the first page of the search results. In other words, for each search query, they were given the first five standard search result pages within the first page. Students were asked to consider all the result links or abstracts in the page (without clicking on any), take mental or written notes of their favourite results and only after reviewing all the search results, click on the links they found most interesting or relevant. After finishing each task, the researcher noted where in the search results students found their desirable information (position of the clicked search result links within the page), and students rated their experience based on timeliness of the search process and relevance of the search results. Evaluation was done based on a Likert scale of 1 (strongly disagree) to 5 (strongly agree).

The study sessions were run using only the Google search engine, following the results in Part I. Measurable Google search personalisation accrues while users are searching with their logged in accounts, assuming these accounts normally store third-party cookies. Personalisation is mostly based on the users' geographical locations, and their account information such as search and browsing history (Hannak *et al.*, 2013).

Students logged into their Google accounts on the search engine's main page to perform the tasks regarding personalised search while their geographical location (Adelaide, Australia) was visible to the search engine. Account cookies and search and browser history were stored normally as set by default settings before and throughout the experiment. To access non-personalised search results, all the following attributes were concealed from the search engine: the student's geological location and IP address; third-party cookies; search and browsing history; and the students' account and demographic information.

The Startpage search engine on a Tor browser was used for the personalised searches. Startpage, an award winning privacy-enhanced search engine, returns the best Google search results without revealing or storing any of user's personal information such as IP address (location), identifying cookies or search histories (EuroPriSe, 2008; Startpage, 2015). During non-personalised search tasks not only were users not logged in (the option is not available on Startpage), but also their geographical location, in case it was leaked to the search engine, was set to countries other than Australia. Third-party cookies and Web history are not stored

in this browser (Tor, 2015). After completing every non-personalised search, the Tor browser session was terminated, automatically clearing all privacy-sensitive data. Mozilla Firefox (version 32.0.3) was used for both personalised and non-personalised search experiments. The browser was modified to accommodate Tor for non-personalised search. Since variations in Web browsers do not result in measurable difference in the search results (Hannak *et al.*, 2013), we consider the effect of this modification to be negligible in our results.

This experiment included five search tasks. The average number of total clicked results per task was fifty-nine summing up to 295 clicked results in total including 165 personalised and 130 non-personalised results. Each student clicked on average two results during each task, after which they filled out a short questionnaire, indicating their satisfaction with the relevance of the information they found, and the amount of time they spent searching for their required information.

The following two sections present the results of this experiment.

Personalised search and information opportunity

Figure 6 shows the click distribution among the five standard pages of the search results, presented to the students in a single page containing all the fifty results, were quite similar during personalised and non-personalised search tasks. This, indeed, is an unexpected result. As personalisation claims to tailor the search results according to each specific user's information needs and interests, it was expected that students find their required information mostly within the first page of the personalised results. This was not the case, as 60% of the students clicked on the results placed not in the first page but in the subsequent personalised result pages. In fact, as it is demonstrated in Figure 6, the first page of non-personalised search results received slightly (4%) more clicks in comparison.

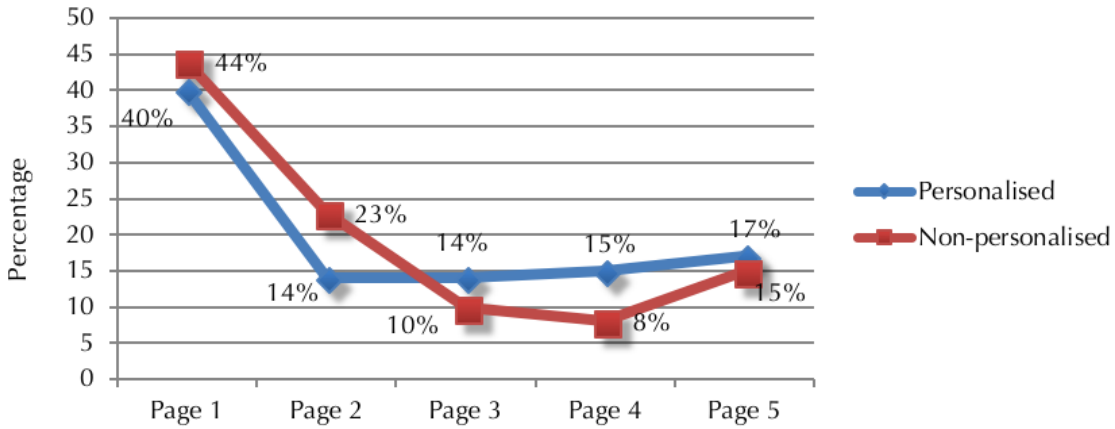

Figure 6: Click rates of personalised vs. non-personalised search result pages

Not only did the first result page of the personalised search contained slightly fewer relevant links than non-personalised search, students also expressed more satisfaction with the quality and relevance of the information they found by non-personalised search (see Figure 7). On average, students expressed complete satisfaction with the information they found during non-personalised search without feeling that it compromised the time efficiency of their information-seeking process. The relevance of information is to a high degree subjective to what users finds relevant to their query and what they need at the time of search. In this experiment the relevance of the search results is solely evaluated from the students' perspective and not by fact-checking all the clicked Websites. According to these results, non-personalised search better met the students' information needs.

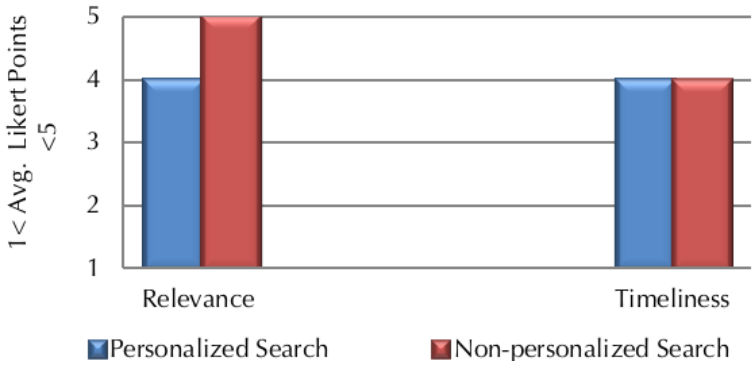

Figure 7: User explicit feedback on the success of the search tasks

Web search and the opportunity bias

This section looks at all the data gathered from this experiment (personalised and non-personalised search tasks) and compares them to Google search statistics, for two reasons. First, because the click contribution within the first five pages of the search results fell into similar patterns for personalised and non-personalised search tasks with only 4% difference in the average number of clicked links on the first result page, and second, because this comparison illustrates how search result ranking and position bias influences students' information opportunity in real-life scenarios.

The objective of this experiment was to see whether, when all the first five standard pages of the search results are given the same visibility to the students, they find the information they are seeking within these pages. This was compared to Google search statistics to see where, within these pages, users normally find their required information. It is already well known that search engine users are biased towards clicking on the first few links within the first page of the search results (Craswell *et al.*, 2008; Joachims *et al.*, 2005). Considering this fact, this experiment clarifies whether search engines, as they state, use this knowledge to place the best information related to the users' queries at the top of the first page or if this dominant bias in users' search behaviour is not being used in their best interest. This section illustrates the relation between search result relevance and its position within the first five standard pages of the search results in controlled and real-life scenarios.

Figure 8 demonstrates what we call the opportunity bias in this paper. Based on students' search behaviour during this experiment, on average 58% of the relevant search result links are not positioned within the first ten results (the first page in the default presentation of search results). By contrast, in real-life scenarios and according to Google statistics, the links positioned on the first page of the search results get almost 90% of the clicks. This indicates missed information opportunities, an opportunity bias, where many relevant search results have little chance of being seen by the users. In other words, users do not have the opportunity to receive information that is relevant to them, because this information is not presented to them in the first page. This experiment shows that users would click on these links once they are seen. However, this does not happen in real-life scenarios as only a few of the relevant search results get the maximum exposure and over half of the relevant links are practically invisible to the users, buried in later pages of the search results.

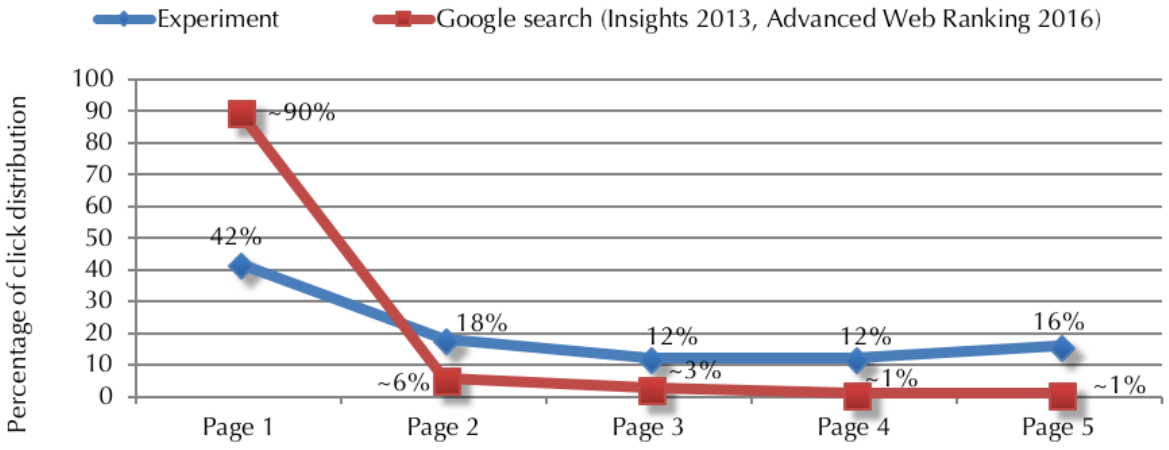

Figure 8: Click rates of search result pages

The difference in search result click distributions between where the most relevant search results are actually located (blue line) and where users tend to click more in real life (red line) could be due to two reasons:

1) The position bias on users' part. Users mostly click on the result links within the first page, as they trust that search engines position the best and most relevant results within the first page. That is why the red line in Figure 8 shows almost 90% of the search result click distribution belongs to page 1.

2) The search engine bias in ranking the search results. Search engine providers are well aware of users' position bias and claim to use this knowledge to better rank the search results: positioning what is more relevant to the users high within the first result page. If this statement was true, the blue line in Figure 8 would follow a similar pattern to the red line, meaning search engines do in fact place the most relevant results within the first page where they are more visible. However, that is not the case. When users are asked to look beyond the first page, they find most of what is relevant and interesting to them in pages 2+ of the search results. This creates the difference in search result click distributions illustrated in Figures 8 and 9. Figure 9 also shows the same data while focusing only on the first result page.

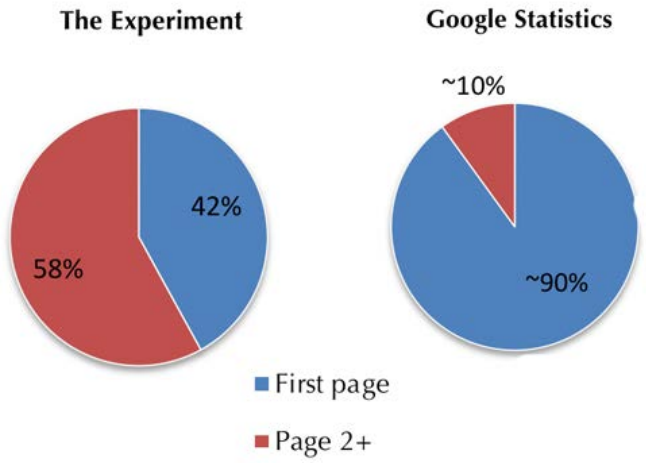

Figure 9: First and subsequent search result pages click rates

As the first page of the search results attracts most clicks, this experiment looks deeper into the positioning of the search results within this page. Research has also shown that Google search engine is overly conservative in changing the link that appears in this position for a query, regardless of the user who submitted that query. This top position is highly coveted due to its exposure and click rates. According to Google statistics, first link in the first result page receives over 30% of the overall search result clicks. As illustrated in Figure 10, that is not the case when users look at more search results.

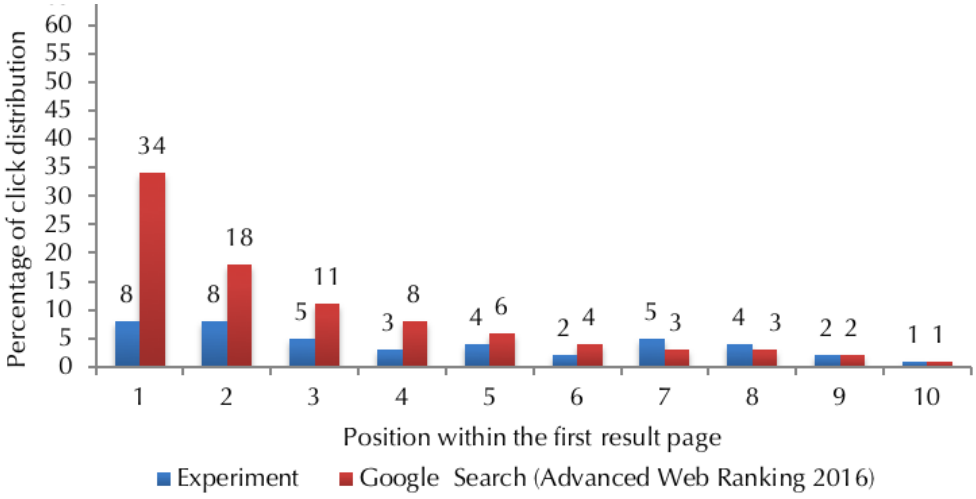

Figure 10: Click rates of positions within the first page of the search result page

In this experiment's controlled environment, the first-ranked search result receives only 8% of the overall click distribution, but also the clicks are distributed significantly more evenly within the first page compared to real-life scenarios. Given that search engine providers declared the main objective of personalisation of search to be better providing users with the most relevant results possible (Horling and Kulick, 2009), the low click rate of the first-ranked search result raises questions about search engines' true incentives in the way they rank the search results. It also points to users' missed information opportunities due to their own behavioural biases in Web search and search engines' exploitation of these biases.

The opportunity bias is not limited to only the positioning of relevant search results, but also to the quality of information. A total of 295 search result links were clicked during this experiment. These links were distributed among .gov (government), .edu (educational institutes), .com (commercial) and .org (non-profit organisations) domains. The .org and .com domains could be purchased by the public whereas the .gov and .edu domains are reserved for official government and educational institutes such as universities. Therefore, the last two domains are generally more prominent in terms of quality and reliability of information (Harvard University, 2017). However, in this experiment 94% of the clicked result links with .edu and .gov domains were positioned within pages 2+ of the search results. In real-life search scenarios, these reliable and relevant links would have had very limited exposure to the users and probably would not be clicked.

Discussion and conclusion

In this study, first we surveyed a group of students from different disciplines, levels of education, age groups and genders about their information-seeking behaviour, specifically, their approach towards commercial search engines and whether they viewed them as a reliable source for research and other educational purposes. The results confirmed that Google is the most popular search engine among students. We also found that not only the majority of undergraduate students, but also doctoral students depend on Google as their primary or even the only source of academic information. More than 83% of students found search engines an important or very important learning resource. Most students stated that not only do they use search engines for educational purposes daily, but also that they also refer to search engines as their first point of enquiry at least once a week. This could mean even if they end up using other information resources, e.g., academic papers, their initial ideas and roadmap to that particular piece information comes from Google search. The degree in which the students relied on search engines for their education was relatively the same for all groups of students.

Moreover, the experiment with twenty-eight university students found that personalisation of the search results has no effect in allocating high position ranks to the relevant search results overall and within the first page. Indeed, students clicked on slightly (4%) more result links within the first page during non-personalised search tasks. They also expressed more satisfaction with the quality of the results from non-personalised search. Additionally, when given equal exposure for the first five standard pages of the search results, 58% of the results that students selected were located in pages 2+ of the search results. This was not at all consistent with what happens in real-life scenarios. According to Google statistics, pages 2+ of the search results receive only around 10% of the clicks. This pattern continues within the first page, where the first-ranked result links receive over 30% of the overall clicks in real life, even though they are likely not the most personalised and relevant search results for the user. This points to a missed information opportunity, where users do not view what they would more likely find relevant.

It seems that Google is providing learning experiences for students and, therefore, has the power to significantly influence society by directing and shaping the mindset of future researchers and professionals. One might argue that it is up to the students to switch to the Google Scholar for their academic needs, as the Google general search engine has not been designed for educational purposes. That is a fair point and students, especially research students, are constantly encouraged to do so. However, discouraging students from using Web search is not a practical solution, as the majority of them still depend greatly, sometimes exclusively, on the Google main search engine for their educational information needs. That is true even when they consider Google as a less reliable source compared to more academic information sources. We need to alter our approach to incorporate education in Web search without compromising search engines' commercial competitiveness or students' information opportunity.

The opportunity bias is an important issue in Web search and education. Higher education students ought not depend on Web search engines that are not designed for higher education-

related purposes. However, this phenomenon and opportunity bias could present an excellent space for education and search engine providers to work towards better understanding students' information needs and ranking the search results that are more desirable to students to within the first result page.

The popularity of Google search engine among students has been growing, despite all the warnings and encouragement they receive to look beyond Google for their education. Understanding and catering to students' particular preferences when it comes to search engines could help:

- The education providers regain control over the flow of academic information, as well as steering students to more reliable and distinguished information resources by positioning them higher within the search result pages where they are more likely to be clicked.
- Search engine providers cater better to one of their most important and loyal clientele (higher education students) while maintaining their commercial interest.
- Students' access the most relevant and reliable academic information through their favourite method and in the easiest way possible.

Taking into account the opportunity bias to alter education and Web search policies is beyond what can be done in the short term or by individual researchers. Despite this, unveiling the opportunity bias and drawing attention to the excess use of Web search by higher education students is an important step towards a practical change in the way education and Web search interact and influence each other. Ignoring these issues could have serious ethical and social consequences for learners and both education and Web search engine providers while incorporating them could present opportunities towards a better search experience and more flexible and efficient learning.

About the authors

Sara Salehi, PhD, graduated from the school of Information Technology and mathematical sciences, University of South Australia. Her research interests lie in personalisation, and experiments related to human information behaviour, Web search, interactive/cognitive information retrieval, and Web system evaluation. She can be contacted at sr.salehi@gmail.com

Jia Tina Du, PhD, is senior lecturer of information studies in the school of Information Technology and Mathematical Sciences, University of South Australia. Her research interests lie in basic, applied, industry and interdisciplinary studies in library and information science, and information management, including theories, models and experiments related to human information behaviour, Web search, interactive/cognitive information retrieval, and Web system evaluation. She can contacted at Tina.Du@unisa.edu.au

Helen Ashman, PhD, is an associate professor in the school of Information Technology and Mathematical Sciences, University of South Australia. Her research interests lie in intrusion detection, authorship attribution (astroturfing and other social manipulation detection, contract cheating detection), privacy and personalisation. She can be contacted at Helen.Ashman@unisa.edu.au

References

Advanced Web Ranking. (2017). *CTR study*. Retrieved from https://www.advancedwebranking.com/cloud/ctrstudy (Archived by WebCite® at http://www.webcitation.org/6y0GXaI8k)

Ashman, H., Brailsford, T. & Brusilovsky, P. (2009). Personal services: debating the wisdom of personalisation. In Marc Spaniol, Qing Li, Ralf Klamma & Rynson W. H. Lau, (Eds.), *International Conference on Web-based Learning* (pp. 1–11). Berlin: Springer. (Lecture Notes in Computer Science, 5686).

Ashman, H., Brailsford, T., Cristea, A. I., Sheng, Q. Z., Stewart, C., Toms, E. G. & Wade, V. (2014). The ethical and social implications of personalization technologies for e-learning. *Information & Management, 51*(6), 819–832.

Brabazon, T. (2016). *The university of Google: education in the (post) information age*. London and New York, NY: Routledge.

Craswell, N., Zoeter, O., Taylor, M. & Ramsey, B. (2008). An experimental comparison of click position-bias models. In *Proceedings of the 2008 International Conference on Web Search and Data Mining* (pp. 87–94). New York, NY: ACM.

Du, J. T. & Evans, N. (2011). Academic users' information searching on research topics: Characteristics of research tasks and search strategies. *Journal of Academic Librarianship, 37*(4), 299–306.

EuroPriSe. (2008). *First European privacy seal awarded*. Retrieved from https://www.startpage.com/uk/press/europrise.html?hmb=1 (Archived by

WebCite® at http://www.webcitation.org/6y0GsB2Gf)

Guan, Z. & Cutrell, E. (2007). An eye tracking study of the effect of target rank on web search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 417–420). New York, NY: ACM.

Haglund, L. & Olsson, P. (2008). The impact on university libraries of changes in information behavior among academic researchers: a multiple case study. *Journal of Academic Librarianship, 34*(1), 52–59.

Hannak, A., Sapiezynski, P., Molavi Kakhki, A., Krishnamurthy, B., Lazer, D., Mislove, A. & Wilson, C. (2013). Measuring personalization of web search. In *Proceedings of the 22nd International Conference on World Wide Web* (pp. 527–538). New York, NY: ACM.

Harvard University. (2017). *Evaluating sources.* Retrieved from https://usingsources.fas.harvard.edu/evaluating-sources (Archived by WebCite® at http://www.webcitation.org/6y0H1F5dy)

Hinman, L. M. (2008). Searching ethics: The role of search engines in the construction and distribution of knowledge. In A. Spink and M. Zimmer (eds.), *Web search* (pp. 67–76). Berlin: Springer.

Horling, B., & Kulick, M. (2009, December 4). *Personalised search for everyone.* [Web log entry] Retrieved from https://googleblog.blogspot.com.au/2009/12/personalized-search-for-everyone.html (Archived by WebCite® at http://www.webcitation.org/6y0HB0htq)

Joachims, T., Granka, L., Pan, B., Hembrooke, H. & Gay, G. (2005). Accurately interpreting clickthrough data as implicit feedback. In *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 154–161). New York, NY: ACM.

Judd, T. & Kennedy, G. (2011). Expediency based practice? Medical students' reliance on Google and Wikipedia for biomedical inquiries. *British Journal of Educational Technology, 42*(2), 351–360.

McClure, R. & Clink, K. (2009). How do you know that?: An investigation of student research practices in the digital age. *portal: Libraries and the Academy, 9*(1), 115–132.

Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G. & Granka, L. (2007). In Google we trust: Users' decisions on rank, position, and relevance. Journal of Computer Mediated Communication, 12(3), 801–823.

Pariser, E. (2011). *The filter bubble: what the Internet is hiding from you.* New York, NY: Penguin Press.

Purdy, J. P. (2012). Why first-year college students select online research resources as their favorite. *First Monday, 17*(9). Retrieved from http://firstmonday.org/article/view/4088/3289 (Archived by WebCite® at http://www.webcitation.org/6y0HIXad4)

Salehi, S, Du, JT, Ashman, HL. (2015). Examining personalization in academic Web search. In *Proceedings of the 26th ACM Conference on hypertext & social media.* New York, NY: ACM.

Sparrow, B., Liu, J. & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science, 333*(6043), 776–778.

StartPage. (2015). *Our privacy policy.* Retrieved from https://www.startpage.com/eng/privacy-policy.html (Archived by WebCite® at http://www.webcitation.org/6y0HQzBa7)

Statista. (2016). *Worldwide desktop market share of leading search engines from January 2010 to October 2017.* Retrieved from https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/ (Archived by WebCite® at http://www.webcitation.org/6y0HYJFIo)

Tor. (2015). *Tor: overview*. Retrieved from https://www.torproject.org/about/overview.html.en (Archived by WebCite® at http://www.webcitation.org/6y0HlZ9t4)

van Deursen, A. J. A. M. & van Diepen, S. (2013). Information and strategic Internet skills of secondary students: a performance test. *Computers & Education, 63*, 218–226.

Walraven, A., Brand-Gruwel, S. & Boshuizen, H. P. (2008). Information-problem solving: A review of problems students encounter and instructional solutions. *Computers in Human Behavior, 24*(3), 623–648.

White, R. (2013). Beliefs and biases in web search. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 3–12). New York, NY: ACM.

How to cite this paper

Salehi, S., Du, J.T. & Ashman, H. (2018). Use of Web search engines and personalisation in information searching for educational purposes. *Information Research, 23*(2), paper 788.

Last updated: 28 May, 2018