Evidence Based Library and Information Practice

Evidence Summary

Study Describes Research Scientists' Information Seeking Behaviour, but Methodological Issues Make Usefulness as Evidence Debatable

A Review of:

Hemminger, B.M., Lu, D., Vaughan, K.T.L., & Adams, S. J. (2007). Information seeking behavior of academic scientists. *Journal of the American Society for Information Science & Technology, 58*(14), 2205-2225.

Reviewed by:
Scott Marsalis
Social Sciences Librarian, University of Minnesota Libraries
Minneapolis, Minnesota, United States of America
Email: marsa001@umn.edu

Received: 17 Sept. 2009 **Accepted:** 04 Jan. 2010

Abstract

Objective – To quantify the transition to electronic communication in information-seeking behaviour of academic scientists.

Design – Census survey.

Setting – University of North Carolina at Chapel Hill, a large public research university.

Subjects – Nine hundred two faculty, research staff, and graduate students involved in research in basic or medical science departments. Participants self-selected (26%) from 3523 recruited. The sample reflected the larger population in terms of gender, age, university position, and department.

Methods – The authors developed a web-based survey and delivered it via PHP Survey Tool. They developed the questions to parallel similar earlier studies to allow for comparative analysis. The survey included 28 main questions with some questions including further follow-up questions depending on the initial answer. The instrument included three initial questions designed to reveal the participant's place and role in the university, and further coding classified participants' department as either basic or medical science. The questions included categorical, continuous, and open-ended types. While most questions focused on the scientists' information seeking behaviour, the three final open-ended questions asked about their opinions of the library and ideal searching environment. Answers were transferred into a

MySQL database, then imported into SAS to generate simple descriptive statistics.

Main Results – Participants reported easy access to online resources, and a strong preference for conducting research online, even when access to a physical library is convenient. Infrequent visits to the library predominantly took place to utilize materials not available online, although the third most common answer for visiting was to take advantage of the library building as a quiet reading space (14%). Additional questions revealed both type and specifics of most popular sources for research, preferred journals, current awareness tools, reasons for choice of journal for publication, and use of bibliographic management tools.

Conclusion – Scientists prefer online tools for conducting library research, although specific contexts influence the preference, and online articles may be printed out for reading or annotation. The participants are taking advantage of the developing online arena, utilizing databases for research, as well as literature searching, access to journals and conference proceedings, and to keep abreast of current research.

Commentary

This study sought to describe the way scientists at a large research university interact with published information. Compared to earlier studies, the authors discovered an increased reliance on electronic publications and a shift to use of the library as place, rather than depository. In seeking to provide evidence to back-up anecdotal trends, the study is to be lauded. However, the value of this study as evidence depends on one's willingness to accept self-reported usage data as valid.

In many ways the study is well executed. The population sample is relatively large, and reflects the demographics of the total population. Appropriate ethics approval was received, and the instrument is included so that others can use it for further work. The description of methodology and report of the results are detailed and well written. However, there are problems with the methodology such that it cannot be recommended for use as evidence. Others may feel that, while imperfect, the methodology is standard enough to excuse its shortcomings.

The most problematic issue with the study design is that the data are self-reported, and largely retrospective. Studies have documented that self-reported usage data are highly susceptible to recall error (Clarke, Fiebig, & Gerdtham, 2008; Pransky et al., 2006). It may have been possible to use resource usage reports, along with IP address ranges or other identifying information to validate the accuracy of self-reported usage; without such validation, the results are questionable. It is probably impossible to test the accuracy of reports of elements such as times the subject visited a library in the past 12 months, the frequency with which they use personal correspondence as a resource, or the percentage of articles in one's collection on which one has made annotations. Thus the study may best be described as reporting scientists' self-perceived behaviour, rather than actual use. A validating study comparing actual use to self-perceptions, when possible, would have strengthened the study, and should be done before utilizing the survey in additional studies. This is not a problem unique to this study; recall data studies are common in the social sciences and have been used in previous studies of information seeking behaviour of scientists (Brown, 1999; Tenopir, King, Edwards & Wu, 2009), however this does not negate the inadequacy of such data for evidence based practice.

Also, the authors do not report whether they used a chi-square test, nor even state confidence intervals, to test for statistical significance. This should have been caught in peer-review, and it calls into question some analysis where responses are close. Although the authors frequently use variants of the word "significant" in their analysis, it is

unclear whether they are reporting statistical significance or personal interpretation.

At the very least the authors should have acknowledged the problem of recall bias, discussed how they balanced this with their choice of study length, and how it influenced their choice of measures of significance.

The survey designers report piloting and refining the instrument over an eight-month period with input from science librarians at their institution. However, at times the questions seem fraught with jargon or librarianship-specific details that may be unreliably understood by the subjects. Do research scientists accurately discern, much less remember, whether the source of an article was an author's Web site or an institutional repository? In responding to question 14, do they consistently classify Google Scholar as a bibliographic database or general web search engine? At times choices are poorly delineated, as when listing two choices for the question "If you visit the library, what are your reasons for going?" as "Use Computers" and "Perform Searches". Thus, while many portions of the instrument seem well designed, others lack face validity.

The authors promised further publications based on the data collected, with more analysis of correlation and comparisons with other studies. Unfortunately these don't appear to have been produced, as the more detailed analysis might have proven more useful. Qualitative or mixed-methods studies could also be highly useful to further illuminate actual behaviours, and validate the findings of this study.

The instrument developed has reportedly garnered a good deal of interest (Hemminger, Lu, Vaughan, & Adams, 2007), and forms the basis of the National Survey of Information Seeking Behavior of Scientists led by Dr. Hemminger's research group. Although no further journal publications appear to have yet come out of the project, I was able to find what appears to be a conference paper using the instrument in a multi-institutional setting (Niu & Hemminger) This research is highly interesting, and has strong potential to inform decision making, however without further validating studies, the instrument and studies using it remain questionable.

References

Brown, C.M. (1999). Information seeking behavior of scientists in the electronic information age: Astronomers, chemists, mathematicians, and physicists. *Journal of the American Society for Information Science, 50*(10), 929-943

Clarke, P.M., Fiebig, D. G., & Gerdtham, U. G. (2008). Optimal recall length in survey design. *Journal of Health Economics, 27*(5), 1275-1284.

Glynn, L. (2006). A critical appraisal tool for library and information research. *Library Hi Tech, 24*(3), 387-399.

Niu, X., & Hemminger, B. (n.d.). Information resources used by academic professors of the United States in the electronic age. In *iSchools*. Retrieved from http://nora.lis.uiuc.edu/images/iConferences/FinalVersion_XiNiu.pdf

Pransky, G., Finkelstein, S., Berndt, E., Kyle, M., Mackell, J., & Tortorice, D. (2006). Objective and self-report work performance measures: A comparative analysis. *International Journal of Productivity & Performance Management, 55*(5), 390-399.

Tenopir, C., King, D.W., Edwards, S., & Wu, L. (2009). Electronic journals and changes in scholarly article seeking and reading patterns. *Aslib Proceedings 61*(1), 5-32.