Search Procedures Revisited

Diane Kelly[1] & Ian Ruthven[2]

[1] University of North Carolina, Chapel Hill, USA
[2] University of Strathclyde, Glasgow, United Kingdom

Introduction

In this paper we pay tribute to our friend, colleague and mentor, Professor Peter Ingwersen, by examining one of our favorite of his papers, *Search Procedures in the Library – Analyzed from the Cognitive Point of View* originally published in Journal of Documentation in 1982 [4]. Like many of Peter's articles it is characterized by a strong theoretical basis that drives and informs empirical investigation, and includes thoughtful discussion of previous research in addition to the research findings.

Search Procedures reflects on a series of studies carried out over a four year period in the late 1970s. It was published at an interesting time for Information Retrieval. Written before Information Retrieval became synonymous with online information seeking it focuses on Information Retrieval within Public Libraries, then the major location for everyday information seeking. While many of his contemporaries focused on information seeking in academic or special library settings, Peter chose instead to focus a setting that was visited by a more diverse set of people with a broader range of information needs.

Search Procedures focuses particularly on the role of the librarian as an intermediary for finding information and the techniques used by intermediaries to understand a library patron's information need. However, already around this time Peter was demonstrating the foresight for which he is known: he predicted (prior to the Internet and Web search engines) that Information Retrieval machinery would become a mainstream technology and that end users would be required to learn how to navigate online searches without the assistance of intermediaries. If Information Retrieval was not to become an elite activity, as he described it in [5], then Information Retrieval interfaces would be required to capture something of the intelligent mediation he investigates in *Search Procedures* or Information Retrieval would become 'a kind of gamble.' [5, p472]. Fortunately, Information Retrieval did not become an elite activity but instead has become one of the most important and popular 'inventions' of the 20th century. Today, information search is a normal part of many people's daily routines and millions of searches are performed daily. While typical search engines are capable of some mediation through features such as spell correction and term suggestion, such mediations

are quite rudimentary compared to the kind that Peter studied and are focused primarily on the query and search results, rather than the person and the information need.

In this article we summarize the main arguments of *Search Procedures* and, almost 30 years after it was written, reflect on its continuing value.

Search Procedures

Like many of Peter's articles, *Search Procedures* is informed by the Cognitive View of Information Retrieval. The Cognitive View is based on knowledge structures or individual cognitive models of parts of the world. Peter observed that each individual's image of the world consists of a '*conglomeration of different knowledge structures*' [4, p170]. This observation was to be the basis for his subsequent theory of poly-representation. Peter identified three major knowledge structures pertaining to the library intermediary: (1) structures around the professional library activities, such as knowledge of documents available for access, knowledge of how surrogates are created, knowledge of how to conduct standard search routines; (2) structures that reflect the librarian's conceptual or domain knowledge; and (3) knowledge structures that reflect the librarian's understanding of the library patron's stated information need and problem situation.

The Cognitive View is concerned with how these three knowledge structures can help mediate between the two other important sources of knowledge structures, those of the library patron who requires information and those of the document authors, which are reflected in the material available from within the library. *Search Procedures* investigates how the intermediaries negotiate these knowledge structures.

Employing a variant of the think-aloud protocol, the study investigates the information search procedures of 13 librarians conducting searches on written information requests and 5 non-expert searchers searching on their own information needs. The non-expert searchers conducted their own searches and only consulted with the librarians if they found no relevant material, leading to the negotiations which were studied. Peter uses the term 'search procedures', giving the paper its name, to reflect combinations of search actions that are performed within a problem-solving task as opposed to 'search strategies' which infer some conscious series of actions. The concentration is, therefore, on the unfolding cognitive reasoning involved in the mediation process as well as the behavioral actions that embody such cognitions.

A particular interest in this article was the creation of what Taylor [17] referred to as the 'compromised information need', a representation of the enquirer's information need. As Peter notes, '*the skill of the reference librarian is to work with the enquirer back to the formalized need…possibly even to the conscious need…and then to translate these needs into useful search strategy*' [4, p178]. That is, the process of negotiation is to help turn the enquirer's

information need into a form that can be used to search the available information, given knowledge of how the information has been represented in the formal systems. This *labeling effect*, requiring enquirers to verbalize their information need into a search statement that may not reflect accurately their information need, is still the subject of much debate; see for example the recent work by Nicolaisen [10]. In *Search Procedures*, Peter does, however, take the position that the labeling effect can misrepresent the actual information need and the role of the intermediary should be to elicit the true information need by a carefully structured dialogue. The label, Peter emphasizes, may be well outside the context of the searcher's real need and the role of the intermediary is to find the right context. Thus, we see in *Search Procedures* an early recognition of the importance of context, a persistent theme throughout Peter's work.

Search Procedures notes that there is not one single patron-intermediary dialogue that is appropriate for all situations. It may be the case, for example, that the librarian is a domain and search expert and, in this situation, will take the lead in the dialogue with the enquirer filling in details. This type of dialogue is referred to as *asymmetrical*. Alternatively, the librarian may be an expert in search but have low knowledge of the search domain, in which case the dialogue is likely to be more *symmetrical* between the patron and librarian. Interestingly, Peter observes that in some cases librarians engage patrons in asymmetrical dialogue because they have too much confidence in their own understandings of patrons' information needs, essentially short-circuiting the process. This, in particular, is a danger when an emphasis is put on speed and least effort. Peter also observes that '*a conscious effort to keep the negotiation on equal-footing would improve the user's chances to provide useful insertions*' [4, p182].

Search Procedures shows that librarians use both open and closed questions to actively build a conceptual understanding of the enquirer's need with concepts being introduced, analyzed, retained or deleted until a suitable understanding emerges that can be used to interrogate the documents. This is described as a type of problem-solving. A surprising feature of the negotiations studied was the low use of 'open' questions: questions that start with 'Why, How, Where,' which should lead to useful information about the context of the information need. Peter's analysis points to the strengths and weaknesses of open questions within the mediation approach as studied: the low use of open questions can limit the enquirer's ability to introduce new concepts and important situational information, whereas over-use of open questions can risk overloading the librarian's original understanding of the need with too much information.

Far more common were 'closed' questions, which Peter divides into *normal closed questions* and *leading closed questions*. Normal closed questions lead to yes or no responses, while leading closed questions present the librarian's expectations about the searcher's answer. In symmetrical dialogue, closed questions can either confirm the librarian's initial understanding of the enquirer's information need

or open up more specific or newer search directions. However, in asymmetrical dialogue when the librarian has a false sense of the enquirer's information need, closed questions can make it difficult to redirect the search.

In the study of librarians searching for the written information requests, Peter points to two search *motives*; what he called 'open search motives' which were actions relating to the discovery of information which could be used to enable the search process and 'fixed search processes' which were actions intended to discover documents themselves. Both motives related to *expectations* of whether a document answering the request may exist at all or what form the answer may be present within documents. Interestingly, different librarians engaged in different modes of operation based on these characteristics: some librarians operated an 'open search mode' where they used heuristic processes to expand their cognitive structures and, then, later in the search, moved towards solving the information problems. In this case the librarians were open to learning throughout the information interaction and often found relevant information later in their searches.

Other librarians were characterized by 'fixed and semi-fixed search modes' where the search routines were more algorithmic in nature with fixed modes describing librarians attempting to immediately retrieve documents that provide answers. This infers a fixed expectation about what answer may be appropriate and in what form the answer will appear. In some cases, semi-fixed modes were employed where librarians would move from fixed mode to open mode for parts of a search, to re-orientate a search, and then move back into a fixed mode. Noticeably open search modes resulted in more search concepts being introduced from the process of searching and less reliance on the written information request.

Revisiting the Findings

How does a study of Public Librarians performed in the late 1970s help Information Retrieval in the 2010s? Firstly, we should note that the relevance of the approaches and debates in *Search Procedures* have persisted since it was published – Michel's [9] paper investigating sources of information used by searchers during cognitive processing, Lin et al.'s paper on the value of reference interviews for understanding information needs [8], Kelly and Fu's work on using open questions to elicit better descriptions of searchers' information needs and contexts [6] and most recently, Nicolaisen's paper on compromised information needs [10] – are only four examples of work that have extended the lines of enquiry developed in *Search Procedures*.

Secondly, the work in *Search Procedures* illustrates the important distinction between what a searcher wants, that is, his information need, and what a searcher

says he wants, that is, his search request. Within operational search systems, the consequences of the label effect is often tackled as a *corrective* feature; search engines ask for an initial search request and then offer functions to modify this request to something closer to the original information need using techniques such as relevance feedback or behavior modeling. What we see far less of in current search engine design are good approaches to obtain the best *initial* search request. Modern search engines, in general, support very well what Peter called fixed-mode searching: immediate access to documents that may provide an answer. There are strong reasons for this: a lot of search requests, such as finding a particular website, are naturally fixed-mode searches which do not require complex interactions; it is easier to develop algorithms for such searches; and, perhaps, giving immediate access to documents gives the impression that search success is close at hand.

In *Search Procedures* Peter points to the weaknesses of relying on fixed mode searching. In particular, fixed mode searching can result in relevant results being overlooked, possibly because the searcher has not been sufficiently informed by the search environment to make good relevance choices, and there appears to be a relative over-reliance on the original search request. This latter point is a frequent concern in Interactive IR evaluations where experimental participants often form queries based on written search requests rather than on the information need that lies behind the request.

Peter points to the fixed mode of searching being a mental attitude rather than a criterion for a good search solution. As he notes: '*It is not possible to rank the search modes in order to point out the qualitatively best one. No doubt each serves specific aims—as their dependence on the working domain shows. The main problem seems to be awareness of the search consequences they cause*' [4, p189]. That is, given a search request or information need, there may be a variety of ways in which we can conduct the search, each of which will have consequences in terms of search success or search satisfaction. Search modes are a choice and choices have consequences. Information Retrieval systems have typically not supported a reflective process into the choice or development of search strategies, providing little feedback on alternative search actions or encouraging searchers to consider the quality of their interactions. Or, as Peter concludes '*the user needs more assistance and better consciousness in these tasks*' [4, p189].

Peter further notes that '*For 'intelligent' online assistants employment of some kind of open search mode seems likely to be the most efficient, because it combines heuristic features in the beginning of the search process with more formal solutions later, using the search algorithms built into the IR system in a flexible way*' [4, p189]. This claim is interesting in light of what has occurred in online searching. The majority of search interface design, the 'intelligent' online assistance Peter predicted, has not supported open-searching but fixed-mode searching. The 'intelligence' has taken the form, not of information dialogues, but of complex statistical modeling of interaction to automatically change search procedures or offer limited forms of interactive query support.

Peter's observations about the types of dialogues that occurred – symmetrical and asymmetrical – also demonstrate a problem with the current interactions people have with search engines. By and far, such interactions are asymmetrical, with leading closed questions being presented to searchers in the form of term suggestion, spell correction and other query variants. Peter points out that the problem with asymmetrical negotiation is that "*far too little new information is exchanged because of the mode of questioning applied*" [4, p181]. Leading questions are compared to the type often asked during a patient-doctor interaction, where the doctor does not allow the patient to clarify his answers, but rather looks for complementary information that supports the initial diagnosis. The principle problem with search is that too much faith is put into the initial interpretation of the information need and subsequent interactions are used to force the need to fit this interpretation.

Revisiting the Method

The methods used in *Search Procedures* are also worth revisiting since they demonstrate a patience and carefulness that is often missing from contemporary investigations of information needs and search behaviors. Data were collected in a real-world setting and included the search procedures used by a mixed group: 13 librarians conducting searches on written information requests and 5 patrons searching on their own information needs. Audio recordings of searchers as they thought-aloud formed the principle data for the study. Peter also observed searchers, documenting their behaviors and actions. In addition, searchers engaged in something Peter called *self-confrontation*, where he and the searchers elaborated on the think-aloud recordings by making '*repetitive runs of the recorded tape immediately after recording, adding comments*' [4, p173]. These supplemental activities – observation and self-confrontation – were designed to enhance the accuracy and validity of the think-aloud data and reflect Peter's ever-present concern for capturing a holistic understanding of search behavior. The importance of this micro-level, intensive perspective on search behavior can be lost at a time when studies of massive search log data are common. However, more of these types of studies are needed if current Information Retrieval systems are to be responsive to searchers, the variety needs they bring to systems and the diverse contexts in which these needs arise and are addressed.

Variations of the methods used by Peter were later used in several studies that examined patron-intermediary interactions including studies by Belkin, et al. [1], Kuhlthau, et al. [7], Saracevic, et al. [13], Spink, et al. [15], and Wu and Liu [19]. The method itself was presented in several review articles of methods used in library and information science including Fidel [2], Harter and Hert [3], Nozomi [11], Saracevic [12], and Wang [18].

Conclusion

Recently, we see a *turn* in Information Retrieval and a re-recognition that searches are often complex processes requiring more cognitive support for searchers and more emphasis on understanding information needs. While Peter and other pioneers of interactive IR have claimed all along that queries are often impoverished representations of a searcher's information need, it has taken a while for IR research to catch-up and acknowledge that single queries often do not tell the whole story and very often even tell the wrong story. New developments in intellectual property searching, legal searching, literature based discovery, biomedical searching, and exploratory searching have demonstrated that only supporting 'fixed mode' searching with 'asymmetrical negotiation' is insufficient for complex search tasks and that very often people are trying to do more than find an answer to a routine question or navigate to a popular resource.

Web search engines are Information Retrieval's most visible success story: they are useful, efficient and we struggle to imagine how we coped without them. However, for many search situations we also need Information Retrieval tools that treat us like adults. That is, we need tools for those situations where we know that our information needs are complex and multi-faceted, where we know that we will need to engage in difficult cognitive work and where we do not expect our need to be satisfied within 1 second and with the exact same search results that were presented to the previous searcher who entered a similar query. The strength of *Search Procedures* is it recognizes that complex searching is a norm not an exception and that good design is not necessarily simpler and faster but more integrative, dynamic and symmetrical.

Suchman [14, p316] claimed that '*..interaction between people and computers requires essentially the same interpretative work that characterizes interaction between people*'. If we are interested in designing intelligent, useful tools for complex search problems then we can find guidance in studies of human-human information interaction of the type described by Peter in *Search Procedures*. In this paper, as in so many others by Peter, we see how research conducted from the cognitive perspective can support modern Information Retrieval.

References

1. Belkin, N. J., Brooks, H. M., & Daniels, P. J. Knowledge elicitation using discourse analysis. *International Journal of Man-Machine Studies, 27,* 2, 1987, 127-144.
2. Fidel, R. Qualitative methods in information retrieval research. *Library & Information Science Research, 15,* 3, 1993, 219-247.
3. Harter, S. P. & Hert, C. A. Evaluation of information retrieval systems: Approaches, issues and methods. *Annual Review of Information Science & Technology, 32,* 1997, 3-94.

4. Ingwersen, P. Search Procedures in the Library – Analysed from the Cognitive Point of View. *Journal of Documentation. 38*, 3, 1982, 165-191.

5. Ingwersen, P. A cognitive view of three selected online search facilities. *Online Review. 8*, 5. 1984, 465-492

6. Kelly, D., & Fu, X. Eliciting better information need descriptions from users of information systems. *Information Processing & Management 43*, 1, 2007. 30-46.

7. Kuhlthau, C. C., Spink, A. & Cool, C. Exploration into stages in the information search process in online information retrieval: Communication between users and intermediaries. *Proceedings of the ASIS Annual Meeting, 29*, 1992, 67-71.

8. Lin, J., Wu, P., and Abels, E. Towards Automatic Facet Analysis and Need Negotiation: Lessons from Mediated Search. *ACM Transactions on Information Systems*, 27, 1, Article 6, 2008, 42 pages.

9. Michel, D. A. What is used during cognitive processing in information retrieval and library searching? 11 sources of search information. *Journal of the American Society for Information Science, 45*, 7, 1994, 498-514.

10. Nicolaisen, J. Compromised need and the label effect: An examination of claims and evidence. *Journal of the American Society for Information Science and Technology*, 60,10, 2009, 2004-2009.

11. Nozomi, I. Approaches in the studies of reference process and its integration: Focusing on the studies of reference interviews. *Library & Information Science, 30*, 1992, 43-58.

12. Saracevic, T. Modeling and measuring user-intermediary-computer interaction in online searching: Design of a study. *Proceedings of the ASIS Annual Meeting, 26*, 1989, 75-80.

13. Saracevic, T., Spink, A. & Wu, M-M. Users and intermediaries in information retrieval: What are they talking about? *Proceedings of the Sixth International User Modeling Conference,* 1997, 43-54.

14. Solomon, P. Conversation in information-seeking contexts: A test of an analytical framework. *Library & Information Science Research, 19*, 3, 1997, 217-248.

15. Spink, A., Goodrum, A., & Robins, D. Elicitation behavior during mediated information retrieval. *Information Processing & Management, 34*, 2-3, 1998, 257-273.

16. Suchman, L. A. Representing practice in cognitive science. *Human Studies.* 11, 1988, 305-325.

17. Taylor, R.S. Question-negotiation and information seeking in libraries. *College and Research Libraries, 29*, 3, 1968, 178-194.

18. Wang, P. L. Methodologies and methods for user behavioral research. *Annual Review of Information Science and Technology, 34*, 1999, 53-99.

19. Wu, M. M. & Liu, Y. H. Intermediary's information seeking, inquiring minds and elicitation styles. *Journal of the American Society for Information Science and Technology, 54*, 12, 2003, 1117-1133.

Addresses of congratulating authors:

DIANE KELLY
School of Information & Library Science
University of North Carolina, 100 Manning Hall, CB#3360, Chapel Hill, NC 27599-3360
Email: dianek[at]email.unc.edu

IAN RUTHVEN
Department of Computer and Information Sciences
University of Strathclyde, Glasgow, G1 1XH, United Kingdom
Email: ir[at]cis.strath.ac.uk