

Supporting the Process: Adapting Search Systems to Search Stages

Hugo C. Huurdeman¹ and Jaap Kamps¹

University of Amsterdam, Amsterdam, the Netherlands,
{huurdeman, kamps}@uva.nl

Abstract. Search engines have become indispensable tools for the information-related tasks performed by a wide variety of searchers across the globe, and the information literacy of these search engine users varies widely. The more complex tasks performed using search engines, involving learning and construction, may consist of multiple stages, potentially affecting searchers' feelings, thoughts and actions. However, despite recent advances in personalization and contextualization, current search engines do not support these stages. This conceptual paper discusses the potential impact of search stages on the desired functionality of search systems. First, it looks at process models in the context of information literacy, followed by the support of current search engines for the stages described in these models. Finally, the paper reconciles the information literacy and system perspectives by discussing novel stage-aware search systems.

Keywords: information search and retrieval, information literacy, information behavior, search stages

1 Introduction

Online search engines have become indispensable tools for the information-related tasks performed by a wide variety of searchers across the globe. The information literacy of these search engine users varies widely, and has been defined as the “ability to recognize when information is needed”, and the “ability to locate, evaluate and use effectively the needed information” [1]. Moreover, the complexity of tasks performed with search engines spans a continuum between simple tasks, such as lookup tasks, and complex tasks, involving learning and construction. The more complex tasks in this spectrum may consist of multiple stages. As Kuhlthau [2] has indicated, each stage in the evolving task process can affect searchers' feelings, thoughts and actions. Kuhlthau's model has similarities with information literacy process models, providing guidance to learners and indicate steps required for successful problem solving [3]. Kuhlthau's process approach has had a “considerable impact” on Library and Information Science, but “little impact” on the design of actual information retrieval systems [2]. Even the current, highly advanced online search engines do not necessarily provide support for the stages described in the models. Previous work looked at ways to bridge *macro-level* information seeking models and *micro-level* search system features [4], and at ways to ‘shake-up’ the shallow information search process of novice searchers [5].

In this paper, we take the perspective of information literacy and information seeking behavior (as defined by [6]), and explore how search systems can support the stages described by two common IL process models.

Section 2 of this paper takes the perspective of information literacy, while Section 3 takes a system perspective. Subsequently, Section 4 looks at how to reconcile the two perspectives, followed by the general discussion and conclusion.

2 The Information Literacy Perspective: Process Models And The Conceptual Implications For Search Support

In this section, we discuss the implications of IL process models for the search support of research-based tasks performed by non-expert searchers. Here, we define non-expert searchers as searchers with limited domain knowledge, procedural knowledge and/or IR system knowledge.

2.1 Information Literacy

Since the concept emerged in 1974, a large body of literature has been written about the concept of information literacy [7], the full review of which is beyond the scope of this paper, and which can be found elsewhere (e.g. [7,8]). In 1994, Doyle [9] has defined information literacy as “the ability to assess, evaluate and use information from a variety of sources.” Over the years, as Lloyd [10] suggests, two distinct views on the concept have emerged. On the one hand, information literacy is viewed as a ‘skills-based literacy’ and information literacy is equated with abilities and information skills related to the information seeking process. On the other hand, information literacy is defined as a “complex phenomenon, which acts as a catalyst for learning” (p.36). In this view, IL is embedded in the learning process, and as Kuhlthau [2] indicates, solely focusing on information skills would neglect the essential stages of “reflecting, constructing, and internalizing to learn and understand for one’s self” (p.164).

Various standards for IL have emerged over the years [7,8]. The ACRL introduced the Information Literacy Competency Standards for Higher Education in 2000, which has been one of the mostly cited standards ever since [1]. In subsequent years, reflecting the evolving views on information literacy, and a rapidly changing higher education and information environment, the standards were further reworked, resulting in the ACRL Framework for IL [11]. It consists of six interconnected concepts, which can be flexibly implemented by institutions. *Threshold concepts* have central place in the framework, which are “passageways or portals to enlarged understanding or ways of thinking and practicing within that discipline” [11].

2.2 Process Models

In addition to the various standards and frameworks, different models describe stages involved in information literacy. We describe two concrete examples, Kuhlthau’s ISP model and the Big6 model. Kuhlthau’s ISP model [2] was based on in-depth research and describes the information seeking process. In its turn, it has inspired many IL process models that followed after. The Big6 model [12] describes skills for information

Table 1. Stages in Kuhlthau’s ISP model (adapted from [2])

Stage	Description
1. <i>Initiation</i>	becoming aware of a lack of knowledge or understanding
2. <i>Selection</i>	identifying and selecting general area, topic or problem
3. <i>Exploration</i>	exploring and seeking information on the general topic
4. <i>Formulation</i>	focused perspective is formed, uncertainty is reducing
5. <i>Collection</i>	gathering pertinent information to focused topic
6. <i>Presentation</i>	completing the search, reporting and using results

problem-solving. Despite the numerous developments in information access since their conception, Kuhlthau’s model can still be used to describe information seeking [13], and the Big6 model for problem solving in the age of the Web [14]. For this reason, and due to the potential applicability in different kinds of research-based tasks, we chose to focus on these two models.

Kuhlthau Based on a number of empirical studies, Kuhlthau has constructed the *Information Search Process (ISP)* model. Kuhlthau’s model describes the stages that searchers go through during information seeking in the context of complex and information-intensive tasks [2]. Table 1 highlights the various stages defined in Kuhlthau’s model. Her studies of the information seeking process have revealed “the importance of forming a focused perspective from information gathered to gain a deep understanding of an issue or question” (p.95). Hence, formulating a focus is an essential element of the information seeking process in the context of learning tasks. One of the novel aspects of this model was that it considered the affective states of searchers: users’ feelings, thoughts and actions evolve during the process. The uncertainty gradually changes, leading to episodes of increased uncertainty, and a generally diminishing uncertainty during the process. At strategic moments, for example of heightened uncertainty, mediators may perform interventions. Kuhlthau’s model has been quite influential and important for the development of the research agenda within the information literacy field [10].

The effects of the stages defined in the ISP model on the interaction with search systems are for instance reflected in a gradual evolution of information sought (from general to pertinent) [2]. Vakkari [15], using an adaptation of Kuhlthau’s model, also observed concrete changes in the perceived *relevance* of information items: in the beginning, domain novices have a low ability to differentiate between relevant and irrelevant items, but this ability may increase during the process. Furthermore, the searchers’ abilities to express their information needs increases (including different search tactics, terms and operators).

Big6 As its authors indicate, the Big6 model is one of the most frequently used information literacy process models in education and practice. Eisenberg and Berkowitz [12] describe it as a “general approach to information problem-solving consisting of six logical steps or stages” (p.5). Hence, it can be applied in the context of student learners, but also to professional or personal contexts. Each stage of the model is “necessary for

Table 2. Stages in the Big6 model (adapted from [12])

Stage	Description
1. <i>Task definition</i>	define the problem and information requirements
2. <i>Information seeking strategies</i>	determine the range of sources and evaluate sources
3. <i>Location & access</i>	locate sources and find information within sources
4. <i>Use of information</i>	engage and extract information from a source
5. <i>Synthesis</i>	organize and present information from multiple sources
6. <i>Evaluation</i>	judge product (effectiveness) and process (efficiency)

the successful resolution of an information problem”, but the stages are not necessarily linear [3]. Depending on context, experience and personal styles, the order of the involved steps can be different, as well as the time spent in each stage. Table 2 shows the various stages included in the model. The stages are “a unified set of information and technology skills” [3], which, according to Eisenberg, are essential for a student to master. The main focus, as opposed to traditional library instruction, lies on the broad problem-solving context, not just the specific skills associated with a certain tool; in other words, the “instructions in specifics comes after instruction in the overall information problem-solving process” (p.7). Inspired by Kuhlthau, the authors indicate that the Big6 approach also “accommodates the emotional or affective side of information-problem solving” (p.14). The Big6 model also supports metacognition, meaning that it aims to create an awareness of learner’s mental states and processes [3]. The practical nature of the model means that it can be relatively easily integrated into curricula, and the authors provide various suggestions to use the Big6 model in educational contexts [3].

Summarizing, both models describe information seeking from a *macro* perspective. Kuhlthau’s model describes higher level aspects of information seeking, while the conceptual part of Eisenberg’s model describes the broad problem-solving context. While having a different focus, both models have many similarities and look at the *process*, i.e. the idea that information skills are not “isolated incidents”, but “connected activities” [3]. Providing support at appropriate moments in the information seeking process, but also a reflexive understanding of one’s own process as indicated in the models may be beneficial to the outcomes of learning tasks performed by non-expert searchers. In classroom and library settings, this support may be provided by instructors and mediators. Considering the pivotal role of online search, however, it would also be desirable if information retrieval systems support learners during the different information seeking and problem-solving stages. The question is, however, to what extent this support is actually available in current IR systems.

3 The System Perspective: Search Support for Stages of Complex Tasks

This section takes the perspective of the IR system. We provide a brief overview of the developments of search support, and current limitations in search support for complex tasks.

3.1 Developments in Search Support

Early command-line information retrieval systems in the 1970s were inspired by the dialogues occurring between (library) intermediary and user [16]. These dialogue-based systems would ask a user questions, similar to a reference interview performed by a librarian, and based on the users' answers would ideally retrieve a focused set of results, usually in the form of a number of references. As Ingwersen has argued [17], also various systems in the 1980s and early 1990s explicitly supported "all stages of task performance" (p.137): these "intelligent intermediary systems" were "to act as an intermediary between an end user and the IR mechanism - and perform similar functions as human expert intermediaries used to perform" (p.162). However, research on these intermediary systems gave way to other approaches. Later IR systems became increasingly streamlined, focusing on query formulation and results list inspection, and leave it to the user to perform the task itself. This can still be observed in current search engines, even though drastic changes occurred in the information environment.

Results in modern online search engines such as Google and Bing are increasingly *personalized*. Personalization, in the context of Web search, has been described as "tailoring search results to an individual's interests" [18]. Personalization can be based on explicit preferences of a user, or based on implicit preferences (i.e. detected by the system). Search results may for example be personalized towards a user's context (for instance location and language), or based on previous interactions with a search engine (for example frequently searched topics).

Various authors have expressed the need to extend the support for open-ended tasks in modern information retrieval systems [19]. *Exploratory search* has been defined by as "an information-seeking problem context that is open-ended, persistent, and multifaceted" [19]. For these open-ended tasks, it is not enough to provide just lookup functionality, but also learning and investigation are important [20]. Exploratory search tasks have some similarities with the initial stages of Kuhlthau's model, in which users engage in task initiation, selection and the exploration of a general topic. Hence, as argued in [4], approaches to support exploratory search on the Web may be valuable to support the early stages of Kuhlthau's model as well. In terms of support, exploratory search prototypes have offered ways to rapidly refine queries, to perform advanced filtering (e.g. using facets), to use visualizations and to perform task management [19]. Second, in the context of Human-Computer Interaction, *sensemaking* has been characterized as a combination of information seeking, analysis and synthesis [18]. Sensemaking may occur in complex, information-intensive tasks, for instance carried out by learners, but also by information analysts. The analysis and synthesis steps have similarities with the later stages of Kuhlthau's model [4], but also with the 'Use of information' and 'Synthesis' stages of the Big6 model, since sensemaking encompasses the "iterative process of formulating a conceptual representation" from encountered information [18]. Concrete support for sensemaking in experimental search systems includes features to group and organize information and to take notes.

3.2 Limitations in Search Support

Despite positive evidence for the usefulness of exploratory search and sensemaking features in experimental settings, online search engines still focus on supporting query for-

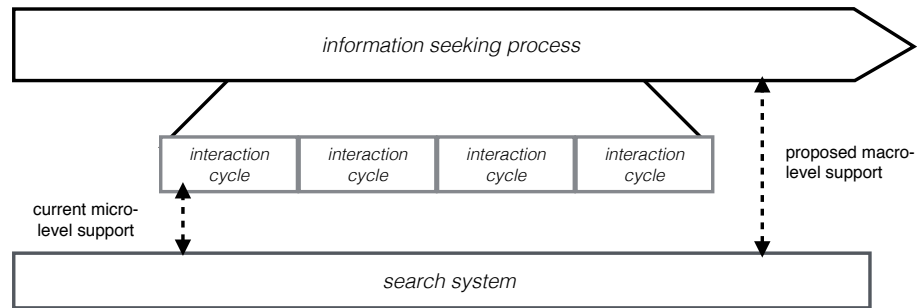


Fig. 1. Micro and macro-level support

mulation and results list inspection, and rarely provide explicit support for complex and open-ended tasks [21]. While search engines' functionality and returned results may be highly relevant to a searcher's query and context, they are not necessarily relevant for the searcher's stage of search. Personalization, for instance, does support displaying search results relevant to individual users' characteristics and preferences, but not the learning or construction occurring within a complex task. In addition, the concept of relevance is pivotal in both information literacy and information seeking. Relevance is multidimensional and dynamic, and is connected to the information needs of a user [cite Borlund]. Retrieved documents, whether relevant or irrelevant, may influence a user's knowledge state and subsequent actions. However, modern search engines do not offer support in judging relevance, or quality of information items.

Summarizing, as Beaulieu [22] has stated, current IR systems may not provide a mode of interaction which is rich enough for task-sharing between user and system. Current IR systems support cycles of micro-level interactions (e.g. consisting of entering queries and selecting results list items), but do not explicitly support the macro-level information seeking or problem-solving stages as described in Kuhlthau's and Eisenberg's models.

4 Reconciling Perspectives: Towards Stage-Aware Systems

Novel ways to support complex tasks performed by non-expert searchers may be inspired by reconciling IL and system perspectives. We look at appropriate ways to increase task-sharing between searcher and system by introducing the concept of adaptive, stage-aware systems.

4.1 Designing Stage-Aware Search Systems

We define a stage-aware system as a potential tool supporting not just *micro*-level cycles of interactions with search systems, but also providing support for *macro*-level information seeking and problem-solving stages (see figure 1). This could for example be achieved by providing different interface panels for each stage, offering customized functionality, content and (guiding) instructions to a searcher.

Stage-Based Adaptation First of all, search tools may be adapted to a user's current information-seeking stage. Search tools supporting information literacy could support the information seeking process by adaptively introducing *functionality* and *content* in a certain information-seeking stage.

Pivotal stages in Kuhlthau's model are 'Exploration' and 'Focus formulation', the latter essentially being a "turning point" in the information seeking process. According to Kuhlthau [2], there are different timepoints in which instructors could intervene, for example at moments of increased uncertainty. In these Zones of Intervention, guidance and assistance may help users to accomplish what they cannot do on their own (p.129). We could extend this view to the search system, and potentially offer different levels of support and assistance in different stages, by means of adapted functionality. For instance, in early stages, with higher levels of uncertainty, more potentially assistive features (e.g. search suggestions) could be included, inspired by features useful for exploratory search. Thus, it could serve to support searchers with a limited domain knowledge, which as Vakkari [15] has stated, "need support to expand and differentiate their conceptual model of the topic", and which have trouble judging the relevance of information items. During later stages, on the other hand, Kuhlthau indicates that users are increasingly able to specify their needs, and to perform comprehensive searches. Hence, in those stages, less support may be needed, but systems could for instance provide functionality for categorizing and organizing encountered results. The design of these features could be inspired by common approaches to support sensemaking.

A second way to adapt search systems to the various information seeking stages is at the *content* level. This could be achieved by selectively showing results, or by customizing the ranking of retrieved search results. From Kuhlthau's model we can derive the importance of showing introductory sources in the early stages, and the idea to not 'overwhelm' the users. This could be performed by ranking introductory sources highly in the beginning of the process, while systems could show specific and in-depth sources (pertinent to the focused topic) more prominently in later stages of the process.

Stage-Based Instruction Both the ISP and Big6 model suggest the positive influence of 'being aware' of one's own information seeking process; for instance by encouraging reflection on the material. A system which aids a user in distinguishing their stage, but also provides search-stage specific guidance, thus may be helpful. Therefore, we can introduce more *prescriptive* search tools, supporting the overarching process and the development of information literacy skills. It is possible to use the Big6 model as an inspiration to design these types of search tools. An integrated tool supporting learning tasks could specifically ask users to define their problem and information requirements ('Task definition'), to determine the appropriate range of sources and weigh criteria ('Information seeking strategies'), and to locate sources and information ('Location and access'), while providing feedback in every step of the process. This can be backed up by literature related to information seeking and retrieval: experimental results by Moraveji et al. have shown that including search tips can have beneficial effects on search skills, even after their experiment finished [23]. The 'Information use' and 'Synthesis' stages have similarities with sensemaking activities, and thus may be supported by system features towards that end, such as notetaking tools. The final stage of the Big6 model is evalua-

tion, encouraging learners to reflect on how ‘effectively’ and ‘efficiently’ a task has been performed. An integrated system supporting the process could use all logged interactions as a way to provide feedback and reflection. Some elements of such a feature can be inspired by Bateman et al.’s ‘search dashboard’ [24], which could be used for “reflection on personal behavior” (e.g. summarizing search techniques and sought topics). The experimental results suggest that users changed search behavior, and their attitudes towards search, based on reflection using the dashboard. The dashboard also provided ways to compare search tactics with ‘expert’ searchers. The latter element corresponds to the Big6 model’s notion that ‘awareness of other styles’ can be helpful. However, despite this positive evidence, to our knowledge, an integration of information literacy stages as described in the process models has not been attempted in systems-oriented research, perhaps due to the various *requirements* of such a system.

4.2 Requirements

An essential requirement of a potential stage-aware system is the detection of stages occurring in a user’s information seeking process. We distinguish between *manual* and *automatic* approaches.

First of all, a system could rely on the input of a user to select which ‘stage’ of an interface to show. This approach has been followed in the *Interactive Social Book Search Track* of the CLEF conference [25]. Searchers for books could manually select panels of an experimental multistage interface, representing exploration, search and review stages in the book search process. The outcomes of this large-scale collaborative study (192 participants) suggest that the users of the multistage interface explore more different kinds of books, and have higher levels of engagement as compared to the baseline interface [26].

Second, a multistage system could rely on automatic approaches to detect a stage a user is in. Considering the complex nature of learning tasks [27], this is not straightforward. To derive the stage a user is in, extensive logging of a user’s interaction with a system is required. For example, evidence could be found in a user’s search terms and tactics, or in her patterns of interaction with a search system [4,28].

5 Discussion and Conclusions

This paper has looked at ways to reconcile information literacy and system perspectives in the context of information seeking. While ‘intelligent’ information retrieval systems from a distant past initiated a dialog with their users to perform task-sharing between user and system, current systems are predominantly focused on queries and results list inspection. However, as discussed in IL literature, non-expert searchers may need additional support during complex and information-intensive tasks, in order to find, assess, evaluate and use appropriate information. These information ventures are inherently dynamic, and the stage a user is in has a profound influence on information sought, judgements of relevance, and searchers’ abilities to express their needs [2]. Hence, increasing the support related to a user’s *process* may have positive effects on the outcomes of learning tasks. This is supported by the positive effects of information literacy interventions (e.g. [29]), but also of experimental search systems encouraging reflection

on encountered materials and search behavior [24,23]. Furthermore, including information literacy instruction into search tools used in the context of research-based tasks may encourage learners to learn by doing, and apply these skills in their later information ventures.

Potential drawbacks of multistage and prescriptive search systems include “lockstep strategies”, which could force “one specific method for problem-solving and decision-making” upon a user [3]. Therefore, stage-aware tools should allow users to flexibly switch between ‘stages’ and interface panels, and a user should be able to remain in control. Also, we have to bear in mind the risks of a ‘tick the box’ approach in the context of information literacy posed by Johnston and Webber: the idea of “reducing a complex set of skills and knowledge to small, discrete units” [30]. This implies a careful balancing of the potential system guidance towards learners.

Our main conclusions are the following. Information literacy, as the countless definitions, models and standards imply, is a wide ranging and evolving concept, of pivotal importance in our current, information-intensive environment. By supporting the synergy of the stages described by various IL process models in actual search tools, we argue that it may be possible to encourage critical use of information, up to the point that it may change searchers’ information behavior.

To increase task-sharing between non-expert user and system, we have introduced the concept of stage-aware tools, which support stages occurring in the information seeking process. We discussed stage-based *adaptation* and stage-based information literacy *instruction*, and pinpointed some of the requirements for stage-aware systems. In ongoing and future work we experiment with these approaches in the context of book search and general Web search [4,5,25,26].

References

1. American Library Association and Association for College and Research Libraries: Information Literacy Competency Standards for Higher Education (2000), <http://arizona.openrepository.com/arizona/handle/10150/105645>
2. Kuhlthau, C.C.: Seeking meaning: a process approach to library and information services. Libraries Unlimited, Westport, Conn. (2004)
3. Eisenberg, M.B.: Information Literacy: Essential Skills for the Information Age. *DESIDOC Journal of Library & Information Technology* 28(2), 39–47 (Mar 2008)
4. Huurdeman, H.C., Kamps, J.: From Multistage Information-seeking Models to Multistage Search Systems. In: *Proceedings IiX*. pp. 145–154. IiX ’14, ACM, New York, NY, USA (2014)
5. Kumpulainen, S., Huurdeman, H.C.: Shaken, not Steered: the Value of Shaking Up the Search Process. In: *Proceedings of the First International Workshop on Supporting Complex Search Tasks co-located with ECIR*. CEUR-WS, Vienna (2015)
6. Wilson, T.D.: Models in information behaviour research. *J. Doc.* 55, 249–270 (1999)
7. Bawden, D.: Information and digital literacies: a review of concepts. *J. Doc.* 57(2), 218–259 (Apr 2001)
8. Virkus, S.: Information literacy in Europe: a literature review. *Information Research* 8(4) (Jul 2003), <http://www.informationr.net/ir/8-4/paper159.html>
9. Doyle, C.S.: *Information Literacy in an Information Society: A Concept for the Information Age*. Information Resources Publications, Syracuse University, Syracuse, NY (Jun 1994)

10. Lloyd, A.: 3 - information literacy practice in academic libraries and the higher education landscape. In: Lloyd, A. (ed.) *Information Literacy Landscapes*, pp. 35 – 70. Chandos Information Professional Series, Chandos Publishing (2010)
11. Association for College and Research Libraries: *Framework for Information Literacy for Higher Education* (2015), <http://www.ala.org/acrl/standards/ilframework>
12. Eisenberg, M.B., Berkowitz, R.: Eisenberg, M. & Berkowitz, R. (1990), *Information problem solving: The Big Six skills approach to library & information skills instruction*. Norwood: Ablex. Ablex, Norwood (1990)
13. Kuhlthau, C., Heinström, J., Todd, R.: *The information search process revisited: is the model still useful?* (2008), <http://www.informationr.net/ir/13-4/paper355.html>
14. Eisenberg, M.B.: *Technology for a purpose: Technology for information problem-solving with the Big6*. *TechTrends* 47(1), 13–17 (Jan 2003)
15. Vakkari, P.: *A theory of the task-based information retrieval process: a summary and generalisation of a longitudinal study*. *J. Doc.* 57(1), 44–60 (Feb 2001)
16. Wilson, M.L.: *Search User Interface Design. Synthesis Lectures on Information Concepts, Retrieval, and Services* 3(3), 1–143 (Nov 2011)
17. Ingwersen, P., Järvelin, K.: *The Turn - Integration of Information Seeking and Retrieval in Context*. Springer, Dordrecht (2005)
18. Hearst, M.: *Search User Interfaces*. Cambridge University Press (2009), <http://searchuserinterfaces.com/book/>
19. White, R.W., Roth, R.A.: *Exploratory Search: Beyond the Query-Response Paradigm. Synthesis Lectures on Information Concepts, Retrieval, and Services*, Morgan & Claypool Publishers (2009)
20. Marchionini, G.: *Exploratory Search: From Finding to Understanding*. *Commun. ACM* 49(4), 41–46 (Apr 2006)
21. Cole, C.: *Google, tear down this wall to exploratory search!* *Bulletin of the ASIST* 40(5), 50–54 (2014)
22. Beaulieu, M.: *Interaction in information searching and retrieval*. *J. Doc.* 56(4), 431–439 (Aug 2000)
23. Moravejji, N., Russell, D., Bien, J., Mease, D.: *Measuring Improvement in User Search Performance Resulting from Optimal Search Tips*. In: *Proceedings SIGIR*. pp. 355–364. SIGIR '11, ACM, New York, NY, USA (2011)
24. Bateman, S., Teevan, J., White, R.W.: *The Search Dashboard: How Reflection and Comparison Impact Search Behavior*. In: *Proceedings CHI*. pp. 1785–1794. CHI '12, ACM, New York, NY, USA (2012)
25. Gäde, M., Hall, M., Huurdeman, H., Kamps, J., Koolen, M., Skov, M., Toms, E., Walsh, D.: *Overview of the INEX 2015 Interactive Social Book Search Track*. In: *CEUR* (forthcoming, 2015)
26. Huurdeman, H., Kamps, J., Koolen, M., Kumpulainen, S.: *The value of multistage search systems for book search*. In: *CEUR* (forthcoming, 2015)
27. Freund, L., O'Brien, H., Kopak, R.: *Getting the big picture: supporting comprehension and learning in search*. In: *Proceedings SAL* (2014)
28. Vakkari, P.: *Exploratory searching as conceptual exploration*. In: *Proceedings HCIR*. pp. 24–27 (2010)
29. Walton, G., Hepworth, M.: *A longitudinal study of changes in learners' cognitive states during and following an information literacy teaching intervention*. *J. Doc.* 67(3), 449–479 (Apr 2011)
30. Johnston, B., Webber, S.: *Information Literacy in Higher Education: A review and case study*. *Studies in Higher Education* 28(3), 335–352 (Aug 2003)