December 1998

Expert-Novice Differences in Searching the Web

Diana Tabatabai
McGill University

Francesca Luconi
McGill University

Expert-Novice Differences in Searching the Web

Diana Tabatabai
Francesca Luconi
Department of Educational and Counselling Psychology
McGill University

Abstract

Three web-user experts and three novices were observed in a problem-solving context. Their task was to search and evaluate information on the World Wide Web in order to write a research paper about an unfamiliar topic. A cognitive task analysis and verbal protocols were used. Results indicate differences in the ways experts approach and solve problem. These differences were in terms of (a) knowledge base; (b) problem space; (c) strategies; and (d) affect. Novice-expert differences have instructional implications for teachers, Information Technologists and Web site designers.

The Problem

The purpose of this exploratory study was to compare three experts and three novices in a web-based problem-solving context. This study differs from other studies on problem solving because the task is open-ended and it also includes problem-finding (Getzels & Csikszentmihalyi, 1975). Our interest in the Web stems from the fact that it has the potential to be a powerful educational tool in that it has unlimited resources and ideas of multiple representations (Owston, 1997; Windschitl, 1998). However, searching for information on the Web can be problematic due to the vastness of its database. Immediate access to a large quantity of information that has not been systematically evaluated by publishers or other authorities might present more complex problems than conducting a search in a traditional library (Fung, 1997). Therefore, the users must apply their critical thinking skills and metacognitive and navigational strategies in order to both successfully evaluate the reliability and validity of the sources (Jones, 1997), and to avoid getting lost in cyberspace. How do experts and novices deal with these complexities in searching the Web? Lessons learned from Web experts have significant instructional implications for educators and designers.

Expertise: Expert-novice differences have been studied to determine the initial state of the learner (novice), and what is required to become an expert. From this research we know that the course of knowledge acquisition proceeds from a declarative to a procedural, condition-action form (Glaser & Bassock, 1989). This difference is between knowing "what " and knowing "how". Novices may acquire the same amount of information as experts without knowing the "condition" of its application. Experts, on the other hand, seem to acquire knowledge in relation to its function and applicability (Chi and Glaser, 1985). Novices tend to categorize problems by surface features whereas experts look for patterns and analytic strategies (Chi, Feltovich and Glaser, 1981). Experts have more in-depth prior knowledge and, through chunking, can switch faster from one strategy to another when faced with difficulty on a problem (Presley and McCormick, 1995). Ericsson, Krampe, & Tech-Romer (1993) believe that it takes 10,000 hours or 10 years in the field to become an expert; but due to the relative novelty of the Web, in this study, we define expertise in terms of hours of Web access and not necessarily years of familiarity with the Web.

Methodology

Our sample consisted of three novices and three experts, all graduate students at the Faculty of Education, McGill University except novice 3 who was a high school student. Experts were Web-based course designers and teachers. Novices, except for 3, shared a common health education and research assistanship background. The criteria that differentiated them was the amount of time spent on the Web. Our novices' time spent on the Web was on average less than 3 hours a week whereas for experts it averaged 15 hours a week. The task was to search information on the Web about an unfamiliar topic (inquiry-based instruction or IBI) for the purpose of writing a research paper. Performance measures were the differences between how experts and novices search the web to solve the problem at hand (see Table 1). Due to the ill-defined nature of the problem, cognitive task analysis and verbal protocols were used to gather information (Shute, 1997; Ericsson & Simon, 1993). The task lasted approximately 1.5 hours. One researcher asked the informants about their rationale for decision making while searching the Web, the other researcher videotaped the sessions. This division of task was kept constant with all participants. To save time, they were given a socio-demographic and training questionnaire to be filled out before the task. After the task was completed, they were asked about different topics related to Web searching. Their responses were triangulated with the results of the questionnaire and the cognitive task analysis in order to verify whether participants' actions during the search coincided with the strategies they told us during the post-task interview. Two coding schemes were created. One was used to analyze data gathered through the PARI methodology (Lesgold, Lajoie, Logan, & Eggan, 1990) that resulted in categories of Table 1. The other was used to graphically represent the participants' problem spaces.

Results and Discussion

Differences between the experts and the novices were analyzed in terms of: (a) knowledge base; (b) problem space; (c) strategies; and (d) affect.

(a) In this study, the knowledge base included prior knowledge of IBI, the Web itself, and the method of source evaluation. None of the participants were familiar with the IBI topic, but the experts knew more about the Web and source evaluation than did the novices. This is reflected in experts using, on average, more key words and combinations (see Table 1.) Another difference was the order in which the Web search and traditional sources such as CD-ROM and Peruse would be used. For example, expert A used the Web to do a coarse grain analysis of what IBI was and to get a feeling of the type of people, organizations, and institutions involved, and would only then go to the CD-ROM for a more specific, fine-grained analysis. Novices reversed this order. In evaluating the sources, the criteria that experts used were authorship, accuracy of content, and currency. Novices 1 and 2 knew about these criteria, probably due to their experience as graduate students, but they did not apply it as systematically as did experts. Novice 3 evaluated the sources based mainly on surface features such as title, organization and content, whereas experts started selecting sources by reading the URL and discriminated sources guided by their prior knowledge about the Web and its jargon, e.g., acronyms such as org. (organization) or com. (company). Based on experts prior knowledge, they could make a better selection of sources as it is reflected in their higher number of relevant sites (see Table 1).

(b) Problem Space is a mental representation of the sequence of actions and decisions that can transform the initial state into a final state or solution (Newel & Simon, 1972; Chen, 1995). We have graphically represented each participant's problem space and created a coding scheme accordingly. Experts had a different problem space than novices. They started the search with a plan whereas novices did not articulate a plan and most of their strategies were trial-and-error. This lack of planning is reflected in the noviceslarger number of return moves compared to experts. The following differences were observed in their problem space: 1) novices had more backward moves; 2) expert used more search engines and sometimes accessed them directly by typing their URL; 3) novices missed high quality sources and were generally not as satisfied as the experts with the search results (Table 1).

(c) Experts and novices differed in the use of navigation and metacognitive strategies (self-regulation and planning). Experts had a more in-depth understanding of navigating strategies. This is reflected in: a) the fewer number of return moves; b) the higher number of search engine; c) the direct access to search engines; and d) the quality of sources found (see Table 1). From a metacognitive point of view, experts reflected on different aspects of the task such as their plan accomplishment and which actions to take next. They started the search with a plan, i.e., starting with a broad search in order to find a description of IBI. They also differed from novices in the ability for setting goals during the search without putting a time limit to their task. Our novices' reflection was mostly self-criticism for lacking the necessary tools to accomplish the task. They did not have a plan and most of their strategies were trial-and-error. Novices 1 and 3 never used the *find* feature to search the key word inquiry in the text. Novice 3 did not know the *bookmark* strategy. However, novices were aware of certain basic navigating strategies; for instance, 2 of the novices kept the three search words together to avoid getting millions of unrelated hits.

(d) In terms of the affect, which included attitudes and feelings, it was observed that despite a lack of familiarity with the topic to search, experts' attitudes were more relaxed, confident, and satisfied with their web search results (see Table 1). The unknown territory made the novices nervous, but for the experts it was business as usual. Experts remained in control of the situation and did not feel physical discomfort whereas novices soon ran out of esteem. Novices became tired and lost due to information overload, whereas experts were not influenced by information overload because they would only look at the first 3 pages of the results, then, after evaluating the relevancy of the sites, decide on a further course of actions. The common thread among all of the novices was a certain degree of frustration due to their lack of knowledge about search engines, methods of narrowing down the results of the search, and specific criteria to judge the validity and reliability of the sites.

In addition to noting expert-novice differences on the cognitive and affective dimensions of Web-searching, this study attempts to raise critical issues on the use of the Web and identifying expert-novice differences that might have significant instructional implications for educators and designers. The implications of our findings are: (a) critical thinking skills have to be taught in order to evaluate multiple sources of the Web; (b) encourage using metaphorical knowledge to map the problem mentally or graphically; (c) in order to avoid reducing attention span and information overload, metacognitive strategics including planning and self-regulation should be taught and modelled not only by teachers but also librarians and peers. Navigational strategies should be taught parallel to critical thinking skills; and (d) exposure to the Web should be done gradually to avoid tiredness and frustration. Identifying personal biases on the use of the Web might prevent negative feelings and attitudes during the search. Although these results were obtained from a small sample and should not be overgeneralized, they are promising and appear to be an important area for further research. Future replications of this study should include larger sample size and inter rater's reliability measures.

References

References available upon request from first author (dtabat@po-box.mcgill.ca).

Table 1. Expert/Novice Differences in Cognitive Task Analysis

Categories	Expert A	Expert B	Expert C	Novice 1	Novice 2	Novice 3
Key words	6	9	4	5	4	3
Combin.	2	6	3	3	3	1
Engines	2	4	3	1	3	1
Relev.sites	9	6	18	2	4	2
Irrelevant	15	26	7	3	5	5
Shared site	0	2	0	2	1	3
Skipped	1	3	1	1	2	0
Tot.sites	26	41	33	7	12	9
Tot.moves	44	56	49	42	48	34
Ret.move	3	3	2	7	7	12
Cur. stop	0	2	0	0	0	0
Satisfac.	++	-+	++	-	+	-

Note. **Key words=**number of words used by the participants to guide their search. **Combin. =**combinations of key words used. **Engines =**number of search engines used during the search. **Relev.sites**=sites judged by the participants as being relevant to the search either verbally or by bookmarking it. **Irrelevant=**sites that were judged by either the participants or the researchers as being irrelevant. **Shared site**=sites that were visited by more than one other participant. **Skipped**=sites that were skipped by the participant but chosen by others. **Tot.sites**=total number of sites visited including search engines. **Tot.moves**=total number of actions including scrolling and using features like *Find*. **Ret.move**=moves to return to the previous position. **Cur. stop**=number of times participants went out on a tangent to bookmark interesting but irrelevant sites for further exploration. **Satisfac.**=the satisfaction of the participants at the end of the search.