Exploration of the Relative Contributions of Domain Knowledge and Search Expertise for Conducting Internet Searches

Eileen Wood, Domenica De Pasquale, Julie Lynn Mueller, Karin Archer, Lucia Zivcakova, Kathleen Walkey & Teena Willoughby

To cite this article: Eileen Wood, Domenica De Pasquale, Julie Lynn Mueller, Karin Archer, Lucia Zivcakova, Kathleen Walkey & Teena Willoughby (2016) Exploration of the Relative Contributions of Domain Knowledge and Search Expertise for Conducting Internet Searches, The Reference Librarian, 57:3, 182-204, DOI: 10.1080/02763877.2015.1122559

To link to this article: https://doi.org/10.1080/02763877.2015.1122559

Published online: 05 Feb 2016.

Article views: 623

View related articles

View Crossmark data

Citing articles: 4 View citing articles
Exploration of the Relative Contributions of Domain Knowledge and Search Expertise for Conducting Internet Searches

Eileen Wood\textsuperscript{a}, Domenica De Pasquale\textsuperscript{a}, Julie Lynn Mueller\textsuperscript{a}, Karin Archer\textsuperscript{a}, Lucia Zivcakova\textsuperscript{a}, Kathleen Walkey\textsuperscript{a}, and Teena Willoughby\textsuperscript{b}

\textsuperscript{a}Wilfrid Laurier University, Waterloo, Ontario, Canada; \textsuperscript{b}Brock University, St. Catharines, Ontario, Canada

\textbf{ABSTRACT}

The relative contributions of expertise in search skills and domain knowledge were examined when using the Internet to find information. Four conditions were compared: expert searchers/high domain knowledge; expert searchers/low domain knowledge; novice searchers/high domain knowledge; and novice searchers/low domain knowledge. Search outcomes and verbal protocols were analyzed. The combination of search expertise and high domain knowledge yielded the most efficient searches. Higher search expertise yielded access to sites rated more accurate and credible. High domain knowledge yielded sites rated more thorough. Verbal protocols depicted searching as a complex decision process. Findings have implications for instructional support.

\textbf{KEYWORDS}

Internet search; search strategies; expert search skills; finding information

The increasing presence of computer technology, especially mobile devices, within classrooms has ensured that the Internet serves as a valuable research and data gathering tool for students at all levels of education (Johnson, Levine, Smith, & Stone, 2010; Mitra & Rana, 2001; Tay, Lim, Nair, & Lim, 2014; Wolbrink & Burns, 2012). Being able to find relevant information from the vast quantities available is a key first step in using the Internet effectively. Growing evidence indicates that the structure and scope of the Internet can overwhelm learners making it difficult for them to harness the benefits of this instructional tool (Dias & Sousa, 1997; Knight & Mercer, 2015; Sun, Ye, & Hsieh, 2014; Wallace, Kupperman, Krajick, & Soloway, 2000). Two cognitive components, domain knowledge and strategic knowledge (Paris & Paris, 2001; Pintrich, 1995; Schunk & Zimmerman, 2012; Zimmerman, 2002), have been identified as key predictors of search success in traditional hard-copy text searches (Downing, Moore, & Brown, 2005; Symons & Pressley, 1993). More recently, domain knowledge has also been identified as an important predictor of effective Internet-based searches (Willoughby, Anderson, Wood, Mueller, & Ross, 2009). The present study extends the existing literature by...
examining the relative importance of domain knowledge and search strategy skill in an online learning task.

**Unique demands of the Internet**

Online searches differ from traditional text searches (Brand-Gruwel, Wopereis, & Walraven, 2009). Unlike traditional sources of information such as textbooks, information on the Internet is not necessarily arranged in a linear format with an introduction, body, and conclusion, nor does it organize or layer information in the hierarchical way that traditional print media organizes information, for example, from most critical to least critical ideas, general to specific information nor through availability to a table of contents, index, or summary—all of which would support and logically guide a learner through the information. In addition, hypertext presents interesting distractions, links, advertisements and other features that can divert learners from both the task at hand and may make discerning critical information more challenging (Zhang & Quintana, 2012). In addition, advertisements, images, links that compete for cognitive resources may result in cognitive overload (Mayer & Moreno, 2002). The Internet is dynamic and constantly changing which means that sites and the content on sites can appear and disappear quickly or be altered substantially over short periods of time which increases challenges in finding information (Dimopoulos & Asimakopoulos, 2010). Although the task of information seeking in any context is a complex problem-solving activity, these additional challenges can make Internet searches more difficult (Henry, 2006; Zhang & Quintana, 2012) especially for learners of differing capabilities (Lidstone & Lucas, 1998) or those not equipped with the information or skills to navigate effectively and efficiently. Understanding the cognitive processes that affect learners’ interactions with the Internet is a critical step toward understanding how to implement this technology effectively, especially in educational contexts (Willoughby et al., 2009).

**Cognitive factors impacting search success**

In traditional text based media, both search strategies and domain knowledge contribute together to produce effective searches from text (Symons & Pressley, 1993) or closed archival information systems (e.g., Psychinfo; Downing et al., 2005). High domain knowledge allows learners to identify and extract key information, organize information for memory, and apply information in higher order educational tasks including problem solving, synthesis, and evaluation. High domain knowledge enhances the efficiency and effectiveness by which learners acquire and use new to-be-learned information (Chi, 1978; Chiesi, Spilich, & Voss, 1979; Hambrick & Engle, 2002; Miller, Stine-Morrow, Kirkorian, & Conroy, 2004; Pressley et al., 1992). When conducting a search through the Internet, higher domain knowledge leads to increased performance,
more efficient searches (i.e., more directed searches), more effective searches (i.e., higher quality of sites searched) (Willoughby et al., 2009), and enables learners to better monitor their learning while using online sources (Moos & Azevedo, 2009). Learners with low domain knowledge generally have fewer resources available to understand and organize information and, in an online environment, this lack of knowledge may impair the ability to guide and understand the interaction with online information (Shapiro, 2004). Because information in traditional texts is arranged in a hierarchical and linear fashion, experienced learners can use their knowledge about text structure to support other search strategies (e.g., locating a particular term from a search of the index or glossary). Having knowledge about texts and their structure as well as domain knowledge about the topic at hand facilitates searches. A lack of domain knowledge may lead to even greater comprehension problems with hypertext than with linear text (Foltz, 1996). The potential learning gains envisioned through use of the Internet, therefore, may only be realized when students have higher domain knowledge.

Learning and performance gains from use of strategies have been demonstrated across a multitude of domains (e.g., reading, mathematics, and sciences) and ages (e.g., Brunstein & Glaser, 2011; Cantrell, Almasi, Carter, Rintamaa, & Madden, 2010; Pintrich & De Groot, 1990; Pressley & Levin, 1983). Having a repertoire of strategies and employing them appropriately facilitates learning. Some strategies provide support when domain knowledge is low, for example, mnemonics strategies support learners’ acquisition of factual content when little or no domain knowledge is available. Higher order strategies (e.g. planning, synthesizing, evaluating, and for Internet searches—advanced searches, Boolean techniques), however, encourage meaningful processing of information and relating to-be-learned information to existing knowledge through elaboration, organization, and problem solving (Pintrich & De Groot, 1990; Pressley & Levin, 1983). In addition to traditional text strategies, there are a number of higher-order strategies that are particularly useful for navigating searches (e.g., Boolean commands). These strategies are often provided as instructional modules through library tutorials or are directly instructed when librarians work with students. Having these higher-order online strategies may be necessary to make challenging tasks such searching for information more efficient and effective.

To date, a limited number of studies have examined the effect of search skill or domain knowledge for finding information online. Willoughby et al. (2009) isolated the relative effect of high versus low domain knowledge and found that University learners with higher domain knowledge displayed increased performance and greater perceived ease in executing Internet searches relative to those with lower domain knowledge. Navarro-Prieto, Scaife, and Rogers (1999) examined web searches with participants who had high and low experience working with the Internet and reported that
those with greater experience planned ahead in their searching behavior on the basis of their knowledge regarding the web. Less experienced users seldom planned and were driven by what was presented on the screen. Lack of planning, limited keywords, and poor evaluation of search results (e.g., reviewing only information provided on the first page of results), are only a few of the ineffective approaches to online searches that have been consistently observed across age and educational contexts (e.g., Knight & Mercer, 2015; Sun et al., 2014; Teevan, Alvarado, Ackerman, & Karger, 2004).

Several studies have noted that searches conducted by novices or younger children tend to be less efficient because of the failure to use advanced search strategies (Carper, 1996; Hall, 2000) and/or fewer or less focused search commands (Zhang & Quintana, 2012). Expert searchers, on the other hand, use more synonyms, a greater variety of search terms and keywords, and they reflect more about the value of sites identified by their searches than novice searchers (Hsieh-Yee, 1993; Teevan, Alvarado, Ackerman, & Karger, 2004). These studies suggest that search strategies may serve important and different functions than domain knowledge for successful searching in online contexts.

One study examined the effects of domain knowledge and strategic knowledge by asking experienced and less experienced Internet users with comparatively lower and higher levels of domain knowledge to search the Internet for information pertaining to the introduction to the Euro currency (Holscher & Strube, 2000). Participants with higher domain knowledge and greater experience using the Internet proved to be most successful in search behaviors compared with their less knowledgeable and experienced peers. This study suggests that the combination of domain knowledge and experience with the Internet conferred an advantage over limitations in both. However, one concern with this study is that experienced searchers were not identified on the basis of educational training or any formal experience with Internet search strategies—they were self-trained, self-reported experts. Furthermore, quality of sites accessed was not assessed. To address these shortcomings, the present study examined both domain knowledge (high vs. low) and strategic knowledge (formally trained vs. no formal training) in a higher education context with participants searching for educational information on the Internet.

**Context for the present study**

The present study investigated the relative contributions of domain knowledge and search strategies for online searches by comparing expert searchers with novice searchers in areas where domain specific knowledge was high or low. In addition, the study allowed an examination of search commands executed, sites located and an assessment of the quality of sites visited (i.e., accuracy, thoroughness, and credibility). Think aloud protocols were employed to understand the decisions learners made as they engaged in the
search activity. These two complementary research methods allowed for a more complete understanding of how domain knowledge and search behavior interact when Internet searches are conducted. On the basis of previous text-based and online search research, three potential outcomes were tested.

1. “Domain knowledge beats all.” It was expected that if domain knowledge is high, strategic knowledge may not be required for effective and efficient searches.

2. “Strategic knowledge beats all.” It was expected that if strategic knowledge is high, domain knowledge may not be required for effective and efficient searches.

3. Both domain knowledge and strategic knowledge are needed in unison for effective and efficient search outcomes.

In addition, the present study examined the process of searching for information among a higher education sample using talk-aloud protocols. These protocols allow a richer understanding of participants’ decisions and actions.

**Method**

**Participants**

Forty participants were assigned to one of four conditions as a function of training in information search skills (high and low) and science domain knowledge (biology; high and low). Reference librarians represented the two high information search skill groups and senior undergraduates represented the two limited search skill groups. Each of the 20 reference librarians had completed a Master of Library Science degree and was employed in one of five university libraries in mid- or large-sized Canadian cities. Ten of the reference librarians (2 male, 8 female) had nonscience academic backgrounds and 10 (3 male, 7 female) had extensive academic training in the sciences. Specifically, among the reference librarians with a science background, all had undergraduate and master of science degrees and 1 had a doctoral degree, and all 10 worked in science libraries, and/or science reference work constituted their primary library responsibilities. Among the nonscience reference librarians, all had undergraduate and master’s degrees in arts programs, and all worked in arts libraries and/or providing general reference support constituted their primary responsibilities. There were no age differences between the two librarian groups, t(17) = -.25, p = .80 (science experience group: M = 45.40, SD = 10.09; nonscience experience group: M = 46.60, SD = 9.50). Librarians were recruited from five universities.

Of the senior undergraduate students, 10 (5 male, 5 female) were registered in a university undergraduate biology program and had completed multiple biology courses (M = 11.50 courses). The remaining 10 (4 male, 6 female)
undergraduate students were enrolled in non-science programs and had never taken any university biology courses. There were no age differences between the science ($M = 22.20$, $SD = 1.30$) and non-science ($M = 21.60$, $SD = 1.00$) undergraduate groups, $t(18) = 1.16$, $p = .26$. All students attended one of two universities located in a mid-sized Canadian city.

Assignment of participant to condition is a key element of the present design. To further confirm differences in the groups as a function of search skill and domain knowledge we examined their self-reported responses to four survey items. Two questions were used to assess each of domain knowledge and search expertise. For domain knowledge, participants rated their level of expertise in biology using a 5-point Likert-type scale ranging from 5 (expert) to 1 (inexperienced) and indicated the number of university biology courses taken in one open-ended question. One 2 (search skill) × 2 (domain knowledge) analysis of variance was conducted for each question and yielded main effects for domain knowledge, $F(1, 36) = 28.65$, $p < .001$, and $F(1, 35) = 39.16$, $p < .001$, respectively, such that those with high domain knowledge reported higher levels of expertise in biology and more biology courses ($M = 2.80$, $SD = 1.11$, and $M = 9.66$, $SD = 6.88$, respectively) than did those with less domain knowledge ($M = 1.45$, $SD = 0.61$, and $M = 0.15$, $SD = 0.49$, respectively). For search expertise, participants rated their perceived search skills using a 5-point Likert-type scale ranging from 5 (expert) to 1 (inexperienced) and indicated the number of advanced courses (college and university) taken where information on search skills was a topic. The resulting 2 (search skills) × 2 (domain knowledge) analyses of variance yielded main effects for search expertise, $F(1, 36) = 13.23$, $p < .002$, and $F(1, 39) = 7.14$, $p < .02$, respectively, such that expert searchers reported having greater search skills and more courses ($M = 4.05$, $SD = 0.51$ and $M = 5.32$, $SD = 3.79$, respectively) than did novice searchers ($M = 3.40$, $SD = 0.60$ and $M = 2.25$, $SD = 3.19$, respectively). These self-report outcomes supported the assignment to condition that was initially made as a function of training. All participants were volunteers and were treated in accordance with the American Psychological Association’s ethical guidelines and each received a small cash remuneration.

**Materials and procedure**

Participants were tested individually at their home university. All participants used one laptop with a 17” monitor. Participants first completed a brief pre-search survey, which assessed demographic information (i.e., age, gender, and highest level of education), and the two questions assessing search skills and domain knowledge used to confirm assignment to conditions. Participants then completed a search practice phase followed by the experimental search phase and, finally a post search survey and interview.
**Search practice phase**
The 20-minute practice phase consisted of two simple Internet search tasks: finding the description of any undergraduate biology course, and finding a website describing the risk factors for stroke. These activities allowed participants to become familiar with the laptop and to use Internet Explorer as their search engine. Participants were also instructed in the talk-aloud protocols (Ericsson & Simon, 1993) and were provided with verbal feedback throughout this practice search time.

**Experimental search phase**
Participants were instructed to conduct one biology search (i.e., “Your task is to search the Internet for information that would be useful in preparing a unit of study in molecular genetics for a grade 12 biology course”). Participants were also given a three point summary of Canadian Ministry of Education guidelines outlining expectations for this task which state that students should be able to (a) “explain the components of molecular genetics, (b) explain processes through laboratory activities and conceptual models, and (c) describe some of the theoretical issues surrounding scientific research into genetic continuity; the general impact and philosophical implications of the knowledge gained; and some of the issues raised by related technological applications.”

All searches and talk-aloud components were audio and video recorded. Tracking software (Track4Win Professional Edition, version 2.1; Sepama Software) recorded all Internet sites visited. Video recordings allowed a full view of the computer monitor. Internet Explorer was opened to www.google.ca at the start of the session. Participants were given 20 minutes to complete their searches.

**Post-search survey**
After the search was complete, participants answered two questions assessing how effective they thought their Internet search was with respect to locating the desired information, rated on a 5-point scale ranging from 1 (not at all effective) to 5 (very effective), and the extent to which they have conducted Internet searches such as this one, rated on a 5-point scale ranging from 1 (never) to 5 (a great deal).

**Interview**
A three-question interview was conducted to assess participants’ experience with the talk-aloud procedure. Participants were asked to identify how they felt while talking aloud during the search task, the experience of the talk-aloud on performance and if they felt their search would have differed if they were not required to talk aloud.
Results

Five sets of analyses were conducted. Three sets examined the sites visited to determine the number of sites participants searched, the quality of the sites accessed, and the amount of time spent on sites. The final two sets of analyses explored the verbal protocols and interview information.

Quantity of sites visited

There was considerable variability in the number of sites visited (range = 21 to 159) with 2,842 sites in total visited across all participants ($M = 71.05$, $SD = 30.67$).

Total number of sites visited

A 2 domain knowledge (high vs. low) × 2 search skill (expert vs. novices) analysis of variance yielded a significant main effect for search expertise, $F(1, 36) = 6.73$, $p < .01$, such that those with greater search expertise conducted fewer searches ($M = 59.90$, $SD = 25.44$) than those with less search expertise ($M = 82.20$, $SD = 31.97$). There was no significant main effect for domain knowledge, $F(1, 36) = 0.40$, $p = .53$. However, the significant main effect of search expertise was qualified by a significant interaction, $F(1, 36) = 6.55$, $p < .02$. Post hoc comparisons indicated that expert searchers with high domain knowledge conducted the fewest searches compared with the other three conditions; no other groups differed (See Table 1 for means).

Examination of sites visited revealed that some participants repeatedly discovered the same site while others accessed new sites. Two additional analyses of variance were conducted to permit examination of the number of repeated and one-time (unique) sites visited (see Table 1 for means).

| Table 1 | Means and standard deviations of total number of sites visited, total number of repeated sites visited, total number of unique sites visited, higher order search strategy, and reading of the material ($N = 40$). |
|---|---|---|---|---|---|
| | Expert search | Low search | Expert search | Low search |
| | High domain | Low domain | High domain | Low domain |
| Total sites visited | $M$ | 46.20 | 73.60 | 90.50 | 73.90 |
| | $SD$ | (18.10) | (24.97) | (31.20) | (32.11) |
| Repeated sites visited | $M$ | 10.10 | 18.50 | 21.00 | 18.60 |
| | $SD$ | (6.45) | (7.28) | (8.86) | (9.88) |
| Unique sites visited | $M$ | 30.30 | 42.80 | 36.50 | 36.50 |
| | $SD$ | (9.3) | (16.10) | (15.74) | (15.74) |
| Higher order search strategy decision category | $M$ | 2.70 | 5.00 | 0.80 | 0.50 |
| | $SD$ | (2.00) | (2.16) | (1.03) | (0.97) |
| Reading of material decision category | $M$ | 0.00 | 0.10 | 1.80 | 6.70 |
| | $SD$ | (0.00) | (0.32) | (1.93) | (3.68) |
Repeated sites visited

Every participant visited a site more than once with an average of 17 repeated sites visited ($M = 17.05, SD = 8.94$; range = 3 to 40). The total number for repeated sites visited was 682 (representing 24% of all sites visited).\(^1\)

The 2 (domain knowledge) × 2 (search skill) analysis of variance yielded a significant main effect, $F(1, 36) = 4.47, p < .04$, for search expertise such that those with greater search expertise visited fewer repeat sites ($M = 19.80, SD = 9.21$) than those with less search expertise ($M = 14.30, SD = 7.96$). There was no significant main effect for domain knowledge, $F(1, 36) = 1.33, p = .26$. The qualifying interaction, $F(1, 36) = 4.31, p < .05$, indicated that expert searchers with high domain knowledge conducted the fewest repeated searches compared with the other three conditions.

Unique sites visited

A total of 1,580 unique sites (range = 12 to 105; $M = 39.50, SD = 17.23$) were visited across all searches, representing 55.6% of all sites visited. This 2 (domain knowledge) × 2 (search skill) analysis of variance also yielded a significant interaction between search expertise and domain knowledge, $F(1, 36) = 5.5, p < .03$. Expert searchers who had high domain knowledge accessed fewer unique sites than expert searchers with low domain knowledge, $t(18) = 2.13, p < .05$, and novice searchers with high domain knowledge, $t(18) = 2.39, p < .03$. No other comparisons were significant.

Quality of sites

Two experts were reviewed and evaluated the content of each site visited. Both experts were senior graduate students enrolled in a university graduate biology program and both had taught undergraduate biology courses. Using 3-point Likert-type scales experts evaluated each site for three quality variables: accuracy ($1 = $ not at all accurate; $3 = $ very accurate), thoroughness ($1 = $ not at all thorough; $3 = $ very thorough), and credibility ($1 = $ not at all credible; $3 = $ very credible). An additional category of not applicable was added to the accuracy categories to reflect sites containing only a table of contents, reference list, or other organizational but non–content-based material. Interrater agreement for 20% of the sites was 95%, 96% and 86%, for accuracy, thoroughness, and credibility, respectively. Disagreements were resolved through discussion. These measures were not correlated.

Three 2 (domain knowledge) × 2 (search skill) analyses of variance were conducted to assess the accuracy, thoroughness and credibility of sites visited as a function of domain knowledge and search skill. For both accuracy and credibility there was a significant main effect for search expertise, $F(1, 36) = 4.90,$
p < .03, and $F(1, 36) = 5.77, p < .02$, respectively, such that those with greater search expertise (accuracy: $M = 2.00, SD = 0.18$; credibility: $M = 2.11, SD = 0.34$) visited sites with higher accuracy and credibility ratings than those with less search expertise (accuracy: $M = 1.86, SD = 0.24$; and credibility: $M = 1.89, SD = 0.22$). There were no other significant main effects or interactions.

For thoroughness, however, there was a significant main effect of domain knowledge, $F(1, 36) = 4.30, p < .05$, such that those with greater domain knowledge ($M = 1.37, SD = 0.26$) visited sites with a higher thoroughness rating than those with less domain knowledge ($M = 1.23, SD = 0.13$). No other main effects or interactions were significant.

**Efficiency in searches**

Searches were scored for strategies consistent with advanced search skills such as flexible use of Boolean search commands to narrow searches (Carper, 1996; Hall, 2000; Wallace, Kupperman, Krajick & Soloway, 2000). Poor search strategies involving the use of repeated commands rather than synonyms or alternate key words were also recorded (Wallace, Kupperman, Krajcik, & Soloway, 2000; Zhang & Quintana, 2012).

**Variety of Boolean techniques**

The 2 (domain knowledge) × 2 (search skill) analysis of variance revealed a main effect of search expertise, $F(1, 36) = 18.18, p < .001$ for use of a variety of Boolean techniques. Expert searchers were five times more likely to use a variety of Boolean search techniques ($M = 1.25, SD = 0.91$) than were novice searchers ($M = 0.25, SD = 0.55$).

**Repeated search commands**

The 2 (domain knowledge) × 2 (search skill) analysis of variance yielded one main effect for search expertise, $F(1, 36) = 9.40, p < .005$, such that novice searchers were more likely to repeat the commands used ($M = 19.15, SD = 15.52$) than were expert searchers ($M = 7.65, SD = 6.19$). No other main effects or interactions were significant.

**Talk-aloud protocol analyses**

Overall, participants complied with talk-aloud instructions. Two raters read all talk-aloud transcripts to extract themes that would capture the thought processes behind the search behaviors. The talk-aloud content reflected a problem-solving task (Brand-Gruwel et al., 2009) marked by a series of decision points. Coding involved identifying each decision point along with the follow-through of this decision. Eleven decision themes/categories were identified (see Table 2 for a full list, definition and example of each). Five of the decisions reflected search skill based decisions (with one category having
Table 2. Summary of the 11 decision category themes identified through the talk-aloud protocols.

<table>
<thead>
<tr>
<th>Label</th>
<th>Definition</th>
<th>Quotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search skill decisions</td>
<td>Search refinement or renewal</td>
<td>A point where participants indicated that a new search or an adjustment of the ongoing search strategy was required.</td>
</tr>
<tr>
<td></td>
<td>Higher order search strategy required</td>
<td>A point where participants clearly identified that a specific higher order search behavior would be needed (e.g., putting terms in quotes, moving to an advanced search, using the “find” function), to narrow or broaden their search.</td>
</tr>
<tr>
<td>Organizing search strategies</td>
<td>Two subcategories indicated that participants were attempting to organize their search strategies. Participants followed through in one of two ways: a) Effectively, by regrouping, consolidating previous findings, accessing/directing the search to a specific site (i.e., “directed”); b) Ineffectively, by simply browsing the related information (i.e., “free-flowing/unorganized”) with no purposeful follow through.</td>
<td>a) “I think that the NCBI site is going to be the best place for this...I’ll just search for NCBI. And once I get to the NCBI site, I’m going to be looking for... they have an education section. b) “I’m just scanning to find a, something”</td>
</tr>
<tr>
<td>Planning</td>
<td>A point where participants stopped their search, followed this pause with a reassessment of their proposed strategies and then followed with a specific goal for next steps.</td>
<td>“I’m going to try to do this step by step, explain the components of molecular genetics...”</td>
</tr>
<tr>
<td>Evaluation using search skills</td>
<td>A point where participants evaluated the material they were accessing in terms of the credibility of the material using previous knowledge of a site, effective, and so forth.</td>
<td>“…I’m pretty familiar with this site from my own personal perspective.” “I wouldn’t want to rely on Wikipedia alone...”</td>
</tr>
<tr>
<td>Domain skill decisions</td>
<td>Evaluations using domain knowledge</td>
<td>Participants commented that they were attempting to determine whether the material was relevant to the topic or otherwise assessed the content using domain knowledge.</td>
</tr>
<tr>
<td></td>
<td>Looking to gain domain knowledge</td>
<td>Participants expressed specifically that what they were attempting to do was to increase their own domain knowledge.</td>
</tr>
<tr>
<td></td>
<td>Acknowledging lack of domain knowledge</td>
<td>Participants expressed concern that their domain knowledge was not sufficient.</td>
</tr>
</tbody>
</table>
two possible outcomes) and three reflected decisions made on the basis of domain knowledge. Search skill decisions reflected efforts to evaluate, refine, or narrow the searches being conducted, and included: decisions to use a higher-order search strategy to enhance the search, organization of the searches, a pause in the search to allow for planning, and an evaluation of the credibility of sites found. Decisions involving domain knowledge reflected evaluations of the information detected, self-directing comments that more personal knowledge would need to be found, and statements reflecting inadequacies in personal knowledge. The remaining two categories captured general decisions involving some evaluation of the process or decisions to simply read material out loud. Interrater agreement with a third rater who independently rated 20% of the protocols was 83%. Disagreements were resolved by discussion. Together, these decision types indicate that searching for information is complex and iterative and draws on search strategies and domain knowledge.

Apart from identifying the kinds of decisions being made, we also examined whether these types of decisions differed as a function of search expertise or domain knowledge. For each of the 11 categories, a 2 knowledge (high vs. low) × 2 search strategies (expert vs. novices) analysis of variance was conducted. Of the 5 search skills decision categories only three yielded significant findings (i.e., higher-order search strategy required, organization of the search strategies being conducted effectively, and evaluations using search skills). Post hoc analyses for the significant interaction for the decision category regarding need for higher-order search strategies, $F(1, 36) = 6.32, p < .01$, indicated that participants high in both search expertise and domain knowledge engaged in more decisions where a higher order search strategy was deemed necessary ($M = 2.70, SD = 2.00$) than were their peers with expert search skills but low domain knowledge ($M = 5.00, SD = 2.16$), $t(18) = 2.47, p < .02$, and both groups of expert searchers engaged these advanced search strategies more often than the groups with limited search experience.

The decision category involving Analysis of the Organization of the Search Strategies yielded one main effect for domain knowledge, $F(1, 36) = 6.19,$

<table>
<thead>
<tr>
<th>Label</th>
<th>Definition</th>
<th>Quotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other decisions</td>
<td>General evaluations</td>
<td>Decision was made about the progress of the search or the material examined but the evaluation was not related to search or domain knowledge</td>
</tr>
<tr>
<td>Reading</td>
<td>Decision to read aloud more than one sentence with no evaluation or other comment to clarify why this material was being read</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. (Continued).
such that those with less domain knowledge (M = 2.40, SD = 1.39) depended on search strategies, such as regrouping, consolidating previous findings, accessing/directing search to a specific site more so than those with high domain knowledge (M = 1.40, SD = 1.14). For evaluations using search skills, a main effect indicated that expert searchers (M = 8.05, SD = 3.61) were more likely to evaluate a site on the basis of strategies such as advance searches, Boolean commands, find function, and credible sites than were nonexpert searchers (M = 3.50, SD = 2.24), F(1, 36) = 22.16, p < .001.

With respect to the three decision categories reflecting domain knowledge, all yielded significant main effects. Specifically, as would be expected, those with high domain knowledge (M = 4.70, SD = 4.38) were more likely to use domain knowledge to evaluate a site than were those with low domain knowledge (M = 0.70, SD = 1.17), F(1, 36) = 15.70, p < .001. Those with low domain knowledge reported having to search for information to increase their domain knowledge during the search task (M = 1.50, SD = 1.36) more often than did those with high domain knowledge (M = 0.55, SD = 0.76), F(1, 36) = 7.47, p < .02. In addition, those with low domain knowledge indicated a lack of domain knowledge to complete the task (M = 2.25, SD = 1.18) more often than those with high domain knowledge (M = 0.25, SD = 0.55), F(1, 36) = 48.98, p < .001.

Among the remaining two decision categories there was one significant main effect for general evaluations such that those with search expertise made more general evaluations of sites (M = 3.50, SD = 2.44) than did those with less search expertise (M = 1.45, SD = 1.05; F(1, 36) = 12.23, p < .002). There was also an interaction for reading indicating that novice searchers with low domain knowledge were more likely to read verbatim information than were all other groups, F(1, 36) = 13.24, p < .002. Novice searchers with high domain knowledge read more than either expert searcher group and expert searchers did not differ from one another. Overall, both search expertise groups rarely engaged in reading. In fact, expert searchers with high domain knowledge never read more than one sentence aloud without adding personal input (see Table 1 for means).

**Post-test survey and interviews**
Participants perceptions about whether or how the talk-aloud protocols affected their search behaviors was assessed both in the post-test survey questions and interviews. Overall, searches were perceived to be moderately effective (M = 3.38, SD = 0.98) and somewhat typical of their normal searches (M = 2.56, SD = 1.47). Because it was possible that search expertise or domain knowledge might influence these ratings differentially, a 2 knowledge (high vs. low) × 2 strategies (expert vs. novices) analysis of variance was conducted for each question. No significant main effects or interactions were detected, largest, F(1, 35) = 2.10, p = .16. Thus, participants did not differ as a
function of search skill or domain knowledge in their perceptions regarding their ability to effectively conduct the searches nor in their assessment of how typical this search was compared with other searches they would normally conduct.

Perceptions about the experience of talk-aloud were explored in more detail through the interview questions. Given that there was considerable overlap in responses to the first two questions, these questions were combined. Two raters used an open-coding technique to capture emerging themes. Reliability was achieved by having a third rater, blind to condition and topic, rate approximately 20% of the data independently resulting in 92% agreement. Disagreements were resolved by discussion.

With respect to participants’ perceptions about what it was like to talk aloud during the Internet search two themes emerged. The first theme involved cognitive processes and the second theme reflected emotional responses to the verbal protocol manipulation. Specifically, the talk-aloud activity was perceived to be effortful which slowed down the search process. Highly automatized behaviors such as skimming were perceived to be more difficult to do when verbalizing aloud (65% reported this) and resulted in less efficient behaviors such as full reading of text (e.g., “I mostly skim very quickly so it was slowing me down a lot...”). Similarly, 55% of participants reported challenges with dual tasking—that is, thinking about what they were doing and thinking about the verbal protocols at the same time (e.g., “I had to think ... and remind myself to talk”; “I was trying to read what I saw on the Internet and what I was thinking at the same time”). Some participants (28%) reported challenges trying to decide what to report and what not to report during the task (e.g., “I wanted to say why I chose some things which I probably wouldn’t have typically said”). It is interesting that the cognitive demands of talk-aloud were identified as familiar demands and nonproblematic by 23% of participants.

There were three distinct affective responses reported by participants: awkward (73%), positive (40%) and changing over time (20%). Participants often reported more than one emotional type of response during the interview. Awkward affect reflected feeling uncomfortable, distracted and embarrassed (“a bit silly”; “kind of weird”; “awkward”). Positive affect was present for those who felt the talk-aloud procedure assisted in the search because it directed and organized the search for them (e.g., “I stopped and had to think what I was going to say at the same time, so I think it helped”; “...it just makes you more aware of what you are looking for...aware of how I was searching”). In many cases participants expressed a change in feeling from a negative starting point to a belief that the talk-aloud task became easier.

When asked whether they thought their Internet search would have been different if they had not been talking aloud approximately half (48%) of the participants indicated that the searches were typical and representative of
everyday searches (“Would have been the same”; “pretty much the same”). However, as noted earlier more than half (55%) indicated that they felt “rushed” and that normally they would take longer and be more thorough. It is interesting that although the search was perceived to be rushed or distracting, the outcome of the search was not expected to have changed as a result of engaging in the talk-aloud. Further, in some cases (13%) the talk-aloud experience was perceived as a useful support: one that directed/focused the search because it served as a task reminder (“helps me work through my thoughts”; “I’m not sure [not talking aloud] would have been as in-depth as it was when I was talking aloud because . . . I was more aware of what I was actually doing”). Finally, 5% of participants indicated that their limited knowledge base affected performance in this search by slowing their progress more so than in typical searches they conduct.

**Searching for structure and structural support**

Two aspects of the data indicated that additional consideration was warranted with respect to ‘structure seeking’. When rating sites for quality the raters identified sites that could not be rated for accuracy, credibility or thoroughness because the vast majority of these sites contained tables of contents, chapter outlines, figures depicting hierarchical structures of information and other similar resources that organize information and provide structure similar to that of traditional hard copy text (i.e., hierarchical and linear) which could not be evaluated for content. A 2 knowledge (high vs. low) × 2 strategies (expert vs. novices) analysis of variance of these sites yielded a significant main effect of domain knowledge, $F(1, 36) = 6.81$, $p < .04$ such that those with low domain knowledge stated searching for structured sites more ($M = 0.32$, $SD = 0.16$) than did those with high domain knowledge ($M = 0.21$, $SD = 0.12$). No other main effect or interaction was significant. In addition, within the talk-aloud protocols, nearly half of all participants (48%) explicitly mentioned that they searched for information that would provide a traditional text structure (i.e., hierarchical and linear fashion), 42% of whom attempted this type of search multiple times.

**Discussion**

The primary purpose of this study was to explore the relative contributions of domain knowledge and strategic knowledge for conducting online searches for educationally relevant material. In addition, the study also provided an opportunity to examine how search decisions are made by learners who differ in search expertise and domain knowledge.
**Search expertise and domain knowledge for searching for information**

Previous studies examining fast-searches online, similar to the searches conducted in the present study, indicate that higher domain knowledge yields better searches as defined in number of sites accessed and quality of sites (Willoughby et al., 2009). The present study suggests that the findings for the role of domain knowledge in the Willoughby et al. (2009) study may have been incomplete. When search expertise was studied in concert with domain knowledge it was the combination of search expertise accompanied by higher domain knowledge that provided the most “economical” searches. Specifically, expert searchers (i.e., reference librarians) with high domain knowledge visited fewer sites and were more likely to visit these sites only one time. This outcome parallels outcomes observed by Holscher and Strube (2000) in their study of experienced versus inexperienced internet searchers. In traditional textbook contexts, learners with high domain knowledge generally are more effective at locating information when compared with their less knowledgeable peers (Chi, 1978; Schneider, Korkel, & Weinert, 1989) and having well developed search skills facilitates learners’ ability to locate relevant information (Pressley & Levin, 1983; Navarro-Prieto et al., 1999; cited in: Fenichel 1979 cited in Hsieh-Yee, 1993; Holscher & Strube, 2000). In text searches, however, domain knowledge typically overrides the need for strategic knowledge. This was not the case in the present study. Instead, having both search skills and domain knowledge offered an advantage in allowing the searcher to access new, unique sites instead of returning to or reaccessing previously viewed sites.

However, the combination of search expertise and domain knowledge was not evident for the quality of searches conducted. Specifically, the accuracy and credibility of sites were greater when participants had higher search strategy expertise while those with higher domain knowledge selected more thorough sites. In terms of quality, expert searchers who would be expected to know how to find information that is relevant to the topic and provided by a reputable source, indeed were able to do so. However, expertise in search strategies did not confer an advantage when it came to accessing sites that could provide complete, detailed and comprehensive information—for that advantage high domain knowledge appeared to have more impact. In the case of quality, or effectiveness, of online searches there appeared to be a trade-off in benefits between search expertise and domain knowledge.

**Comparing skilled and less skilled searchers**

Less skilled searchers—those not formally trained in search strategies—tended to repeat search commands. For example, many of these participants would type “Molecular biology” in a search engine once, if they did not find information they desired they would then type “Molecular biology” for a new search failing to
realize that repeating a command yields the same output. These types of search errors indicate a fundamental failure to understand how searches are performed within search engines. These findings are consistent with previous research findings that advanced searchers use synonyms, and more varied combinations of phrases (Sun, Ye, & Hsieh, 2014); than novices. Similarly, the use of Boolean search techniques was expected among all participants in this study as the age of the participants clearly identifies many as ‘digital natives’ (Prensky, 2001). That is, many participants would have lived all of their lives being surrounded by and having access to computer technologies so some knowledge of Boolean search commands would be expected. However, flexible use of Boolean commands was restricted to the expert searchers. Indeed, expert searchers were five-times more likely to use a variety of Boolean search techniques when compared with novices. These expert searchers were able to recognize and tailor their searches to locate valuable information on the Internet.

Together, the aforementioned findings suggest that some generalized knowledge about search strategies was known by most participants in the study and it would not be unexpected to find similar levels of basic familiarity among most Internet users. However, the errors and ineffective search behaviors evident within this sample suggest that most learners would benefit from explicit instruction regarding strategies that yield efficient and effective searches. Many libraries provide instruction, online modules or fact sheets to encourage users to use these strategies but even these supports may require more explicit instruction and modeling to be acquired. Instructional interventions may need to specify how each search command yields specific outcomes; how simple alterations using synonyms are necessary to yield new outcomes, and an introduction to Boolean terms and how to combine these terms to narrow searches.

**Other factors affecting searches**

It is interesting that nearly half of all participants (48%) searched for information that would provide a traditional text structure such as Tables of Contents for textbooks. A similar proportion of participants repeatedly attempted to access this type of information throughout this short search time. This specific search type may reflect a need to access material presented in a hierarchically and linearly organized manner to provide participants with cues as to whether the information they were locating was or was not “on track,” or whether the information was an important concept or subsidiary idea and whether the information found was sufficiently detailed. Indeed, those with less domain knowledge sought this structural support for organizing information. In addition, many participants indicated a lack of confidence in their knowledge while talking aloud and some reflected that they needed some source of information to reaffirm if they were on the right track.
These findings suggest that online searches may indeed be particularly challenging and may require structural scaffolds to maximize outcomes, especially for less knowledgeable learners. These structural supports are often the very structures used to construct texts and organize information in a meaningful way (Raes, Schellens, De Wever, & Vanderhoven, 2012; Rienties et al., 2012). Alternatively, similar to the strategies demonstrated by the librarians, learners might need instruction on how to search out traditional text structures as part of their online search. Tables of contents, glossaries, chapter summaries, and so forth could be used to confirm whether the searcher is on track. Supporting the development of domain knowledge through highly structured materials while providing explicit strategic instruction for searches, may be the kind of complementary instruction that is needed to maximize learning from the Internet.

**Verbal protocols**

The real-time verbal protocols provided insight with respect to characterizing the cognitive events involved in searching for information online. Understanding searches as a series of decision-making choices provides a framework for the flow of search events. Search strategies, allowed participants to refine or narrow the search to find “better” sites, or reorganize or structure the information found. All but one of the search strategies identified in this study were decisions oriented toward better organization and more comprehensive sites being accessed. However, one search strategy led to general viewing and was not a particularly effective decision. With respect to domain knowledge decisions, participants sought to evaluate the material presented to them in these sites as well as evaluating their own personal knowledge. Often a personal lack of domain knowledge was acknowledged and time was spent increasing domain knowledge through searching for a site that provided a brief overview.

It is interesting that analysis of decisions involving search strategies did not consistently favor participants with search expertise. In two cases decisions involving search strategies favored those with greater search expertise but in one situation decisions involving search strategies were made more frequently by those with high domain knowledge. Specifically, expert searchers were more likely to pause and decide that they needed to change their search and employ more higher-order search strategies than were those with less search skills. In addition, expert searchers with low domain knowledge were more likely to use more specific search strategies—relying on search skill to compensate for lack of domain knowledge. In the one case where domain knowledge was critical for making decisions about search strategies, those with low domain knowledge made decisions favoring search strategies such as regrouping and consolidating previous findings to direct their search.
It is interesting that expert searchers also tended to evaluate their searches more than nonexpert searchers. In general, these outcomes confirm that expertise with search strategies influences search behavior. In addition, when knowledge base fails to facilitate the search task, those with low domain knowledge acknowledge the need for search strategies, but in the absence of high-order strategies they engage in review of their current and previous search behaviors to indicate what to do next.

All decisions involving domain knowledge were reflected by main effects for domain knowledge. Those with high domain knowledge were able—and did—evaluate site content on the basis of their knowledge. When domain knowledge was low, participants were more likely to acknowledge their lack of domain knowledge and demonstrated direct effort to increase that domain knowledge to proceed.

Having limited domain knowledge and little search expertise led to reading information from the Internet verbatim. This was not found for those with high search expertise and high domain knowledge. Reliance on lower-order learning strategies such as rote repetition is not uncommon for learners with limited knowledge (Pressley & Levin, 1983; Pressley et al., 1992). However, for some novice searchers, especially those with low domain knowledge, acquiring information is an important first step to solving more complex tasks. Reading aloud may be an important indicator of a struggling learner which educators could use as a cue for providing additional instruction and scaffolding.

**Understanding the context of the search task with respect to typical searches**

Analyses of the survey and interview responses reflected mixed responses to the talk-aloud procedure and to participants beliefs about their search experiences. In terms of cognitive demands, talking aloud was perceived by some to slow their time on task and engaging in verbal protocols may have required learners to multi-task. Given a wide body of research that indicates that performing two similar tasks (for example, two verbal tasks—reading while talking) can compromise or slow performance (e.g., Pashler, 1994; Wood et al., 2012), use of verbal protocols may have slowed search activities but it did not appear to compromise the searches. It is interesting that when asked in the interview, many participants indicated that their search was not greatly affected by the talk-aloud. Participants indicated that their search behavior would have remained similar, with differences lying primarily in the slower speed in which the search was conducted. Therefore, the verbal protocols do not seem to have affected decision-making processes during the searches. The extension of talk-aloud protocols to online search activities suggests that verbal protocols are also a relevant method for studying cognitive activities in this domain. This adaptation
offers a potential tool through which researchers can study how users navigate the use of technology.

Limitations and future directions

Expert searchers occasionally used their knowledge of, and exclusive access to, various library systems to navigate to sites that could not later be accessed and rated through the software available. Although a trace amount of information was available to allow for an estimation of the quality of the site (an address identifying a source—for example a government document/site) access to these sites would have permitted a better understanding of the relevance and importance of the information contained in these sources. In addition, this indicates another “layer” of search skill beyond online search skills that was unique to the highly trained searchers in the sample. Knowing more about the sites they accessed and how they used the information at these sites could have implications for future interventions aimed at developing advanced search skills.

Online data gathering and activities are a common part of the daily lives of students and many students are encouraged to contact library support to conduct their investigations—and reference librarians clearly have skills that could facilitate student learning. Understanding these critical cognitive underpinnings for online searches provides a mechanism to design instructional support for learners in this unique and challenging educational task.

Notes

1. The scores for repeated and unique sites do not add to 100%. Repeated sites were scored as repeated only once, even though some users visited the same site multiple times.

2. The scores for affect do not add to 100% because scoring on affect was not mutually exclusive; some participants indicated more than one affect.

Acknowledgments

The authors thank all of the librarians and students who agreed to give us their time to complete this study.

Funding

This research was funded through a grant from the Social Sciences and Humanities Research Council of Canada.
References


