Computers & Education 52 (2009) 640–648

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

Fast searching for information on the Internet to use in a learning context: The impact of domain knowledge

Teena Willoughby [a,*], S. Alexandria Anderson [b], Eileen Wood [c], Julie Mueller [c], Craig Ross [c]

[a] Brock University, 500 Glenridge Avenue, St. Catharines, Canada
[b] University of Waterloo, Waterloo, Ontario, Canada
[c] Wilfrid Laurier University, Waterloo, Ontario, Canada

ARTICLE INFO

Article history:
Received 4 May 2008
Received in revised form 5 October 2008
Accepted 2 November 2008

Keywords:
Internet
Domain knowledge
Learning
Search behaviors

ABSTRACT

The purpose of the study was to examine the role of domain knowledge when retrieving and using information from the Internet as a resource for essay tasks, as well as to investigate the quality of Internet searches and its relation to essay performance. In two experiments, 100 undergraduates searched the Internet for 30 min and completed two essays; one in which they had high domain knowledge and one in which domain knowledge was low. Two control groups of 70 undergraduates just wrote the essays. Searching the Internet for information enhanced essay performance relative to the control groups only for the topic for which participants had high domain knowledge. In the second experiment, analyses of Internet searches revealed large individual differences in search behaviors and these behaviors did not relate to essay performance, although individuals highlighted the importance of domain knowledge in making their searches easier. Domain knowledge is one factor that educators should pay attention to when using the Internet for learning tasks, particularly when study time is limited, in order to maximize the ability of students to successfully retrieve and use information from the Internet.

1. Introduction

What factors separate children and adults who are "successful learners" from the children and adults who are "less successful learners?" Within the field of cognitive development, many theoretical and empirical papers identify successful learners as those who engage in self-regulated learning (e.g., Pintrich, 1995; Zimmerman, 1989). In other words, these learners have extensive domain knowledge, are intrinsically motivated to learn, engage in metacognitive behaviors that allow them to monitor their behavior and performance, set goals, use sophisticated strategies, and often coordinate many strategies at once (Perry, VandeKamp, Mercer, & Norby, 2002; Willoughby, Wood, & Khan, 1994; Willoughby, Wood, & Kraftcheck, 2003). Our knowledge about successful learning has been generated primarily from traditional text contexts but many learners are now using the Internet as their main source for information. It is important, therefore, to determine whether these earlier findings generalize to computer learning environments. The present study examined one aspect of self-regulation, that is, the role of domain knowledge when information from the Internet is retrieved and then used in a learning context. Within the context of the Internet in particular, the role of domain knowledge may be particularly salient (Bilal, 2001; Downing, Moore, & Brown, 2005) as the structure of the Internet, for example being nonlinear and dynamic, may create unique challenges for learning that are not found in traditional text learning environments (Kuiper, Volman, & Terwel, 2008; Pritchard & Cartwright, 2004). The first experiment in the present study focused specifically on the role of domain knowledge in an Internet learning context and the second experiment examined the quality of Internet searches and its relation to performance.

2. Background

2.1. Importance of domain knowledge

When children and adults are compared on learning tasks, children typically perform less well than adults. This outcome, however, can be reversed when children are "experts" in a domain and have higher domain knowledge than adults. For example, in a classic study, chil-

* Corresponding author.
E-mail address: twilloug@brocku.ca (T. Willoughby).

0360-1315/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compedu.2008.11.009

dren with expertise in chess recalled chess maneuvers more accurately than adult chess novices (Chi, 1978). Similarly, among adult populations, those with high domain knowledge in baseball recalled more relevant information from a text passage about baseball than adults with low domain knowledge (e.g., Chiesi, Spilich, & Voss, 1979; Voss, Vesonder, & Spilich, 1980). Domain specific knowledge, therefore, is a critical foundation for higher level cognitive processing. In memory tasks, domain specific knowledge influences not only what individuals recall but also how much they recall (Schneider, Korkel, & Weinert, 1990). Domain specific knowledge can facilitate the use and generalization of cognitive strategies. For example, individuals with higher domain knowledge typically have more effective search strategies to find and extract information from text (Symons, MacLatchy-Gaudet, Stone, & Reynolds, 2001; Symons & Pressley, 1993) or closed archival information systems such as Psychinfo (Downing et al., 2005). In fact, if knowledge base is high enough, knowledge can override the need for strategies (Schneider et al., 1990).

Furthermore, domain knowledge, in interaction with the other components of self-regulated learning, serves an important function both for acquiring information and for retrieving that information at a later time. For example, domain knowledge facilitates organization of to-be-learned information and such organization impacts on retrieval. According to cognitive schema theory, our knowledge about concepts is stored within interrelated networks of more general information (Anderson & Pearson, 1984). The ease of access, or search, of this knowledge is thought to depend on the richness and level of elaboration of the network (Rumelhart, 1980; Willoughby et al., 1994). When domain knowledge is low, schemas may be insufficiently developed, making the process of creating connections more challenging, less efficient, and less automatic. As a consequence, fewer working memory resources are available for sophisticated processing of the material. In contrast to experts, therefore, novices may be more likely to be overwhelmed when faced with a challenging or unfamiliar task.

2.2. Searching for information on the Internet

Searching for information on the Internet may be an example of the kind of challenging or unfamiliar task that is likely to separate learners who have high domain knowledge from learners who have low domain knowledge, specifically because of the Internet's structure and organization. The Internet consists of hypermedia, an open-ended structure that some researchers suggest promotes active learning and encourages knowledge acquisition through the availability of multiple representations (e.g., Jacobson & Spiro, 1995; Sewell, 1990). In addition, the Internet allows for the ability to access vast and diversified information. While the potential benefits of using the Internet may be extensive, the sheer volume of information and the lack of constraints on the content present a daunting challenge for novice learners to navigate. Learning through Internet use, therefore, warrants investigation.

In previous research, investigators have argued that students may become overwhelmed with the amount and structure of information available to them on the Internet, especially given its nonlinear and non-hierarchical structure, dynamic nature, and lack of quality control (Dias & Sousa, 1997; Kuiper et al., 2008; Lidstone & Lucas, 1998; Slatin, 1990). That is, unlike traditional sources of information, information on the Internet often is not arranged in a linear format with an introduction, body, and conclusion, which would logically guide the reader through the information. Similarly, the Internet typically does not organize or layer information in the way that traditional sources organize information, for example, from most critical to least critical points, or from general to specific points (i.e., it is non-hierarchical).

Added to this organizational constraint of the Internet is the vulnerability of information. The Internet is dynamic and constantly changing, with sites appearing, changing and disappearing unpredictably. Finally, the Internet is missing a critical editor – no reviewer, agency or governing body is responsible for screening the quality of material and evaluating the content in terms of bias, accuracy (including scientific merit) and accessibility of the material for readers at different levels of literacy. These features of the Internet may impose a higher cognitive demand on the learner and cause difficulties in comprehension (Lidstone & Lucas, 1998). In addition, extensive work by Mayer and colleagues (e.g., Mayer, Heiser, & Lonn, 2001; Mayer & Moreno, 2002) suggests that cognitive overload can occur quickly if learners are faced with information that competes for cognitive resources, as is often the case in a hypermedia learning environment where there are multiple modes of representation on the same site (e.g., text, video, and sound, as well as distracting information, such as pop-up ads).

Although the Internet offers a wealth of opportunities for acquiring information, exposure to new information may not necessarily ensure rich elaborative processing (Harris, 1996). Given the important role of domain knowledge for facilitating memory processes (e.g., Schneider et al., 1990), a lack of domain knowledge may lead to even greater comprehension problems with hypertext than with linear text (Foltz, 1996). When dealing with hypertext, the learner has to deal with the cognitive effort required to navigate hypertext and the learner must expend cognitive effort to understand the unfamiliar information. The potential learning gains envisioned through use of the Internet, therefore, may be more limited when students have low domain knowledge.

In one study examining this issue, Park and Black (2007) found that the level of domain knowledge that the participants possessed had a significant impact on their being able to successfully find information from the Internet to answer fact-based questions; however, the study did not include a control group to assess whether the answers were of better quality than if they had not searched the Internet at all. Desjarlais and Willoughby (2007) assessed differences between novices and experts in using information retrieved from the Internet for an essay-writing task, and included a control group that did not search the Internet, but found that novices did not differ from experts on their essay scores. In this study, participants were given an extensive amount of time to search the Internet. Allowing such an extensive amount of time to search the Internet for information is not always feasible, however, especially in an educational setting, and it is critical to explore whether this finding holds true also when the time available to search the Internet is more limited.

Furthermore, an important objective is to see how Internet searches might differ across knowledge areas. Internet search behaviors are extensive and include how many search engines are used (e.g., MSN, Google), the number of terms used in the searches, how many links are accessed, access to relevant versus irrelevant sites, and so on (e.g., Park & Black, 2007; Tu, Shih, & Tsai, 2008). It could be argued that the significantly higher essay performance associated with high knowledge domains as compared to low knowledge domains might be attributable to more effective and efficient search behaviors. Poorer performance in low knowledge domains, on the other hand, may be the product of disorganized, scattered and less efficient search behaviors in that domain (see Hembrooke, Granka, and Gay (2005) for a study of differential search term selection between experts and novices). Alternatively, performance differences between knowledge domains may not be contingent on search behavior. Consistent with schema theory (Anderson & Pearson, 1984), when individuals encounter new information on the Internet, they are better able to encode, organize, and retrieve that information when they have high domain knowledge. In such a case, domain knowledge rather than search behaviors would be the critical factor. To address these issues, two experiments were conducted in the present study. The first experiment focused on the role of domain knowledge when retrieving and using

information from the Internet as a resource for essay tasks, and the second experiment examined the quality of Internet searches and its relation to performance.

3. Experiment 1 – the role of domain knowledge when retrieving information from the Internet and then using that information in a learning context

In order to examine the role of prior knowledge, it was important to construct a task that was conducive to manipulating high versus low domain knowledge without priming participants for their domain knowledge about the topics prior to study (i.e., by pretesting). To do this, each participant was given two topics to research on the Internet (i.e., one topic about biology and one about urban environment). Participants had high domain knowledge for one topic and low domain knowledge for another topic, that is, participants were drawn from one of two fields of specialization [science (SC) or environmental studies (ES)]. Topics were selected in consultation with program instructors to ensure that participants would have background knowledge in the topic selected for their area of specialization although no specific instruction in the topic. Together, these aspects of the design permitted a within- and between-subjects comparison of performance as a function of knowledge base. Specifically, the within-subjects comparison involved asking participants to search for information and write two essays (one relating to a high knowledge domain and one to a low knowledge domain). The between-subjects comparison allowed us to compare the relative performance of participants from each of the two disciplines (i.e., SC or ES).

In order to test differences in a topic domain area while using computer technology, it is important to control for the effects of Internet search experience. As with content knowledge, knowing how to search the Internet to access information may be a critical factor in the process of learning (Gergets & Hellenthal-Schorr, 2008; Hess, 1999; Kuiper et al., 2008; Marchionini, 1995; Mitchell, Chen, & Macredie, 2005). Without appropriate knowledge of search skills (such as how to use search engines, Boolean search techniques, advanced search strategies, etc.), learners may become overwhelmed by the vast amount of information available to them. To avoid this potential confound, the present study manipulated whether participants did or did not receive access to self-teaching online training in Internet search strategies prior to conducting their search. In order to measure learning, students completed essays after searching for information on the Internet. Finally, in order to determine the value of exposure to the Internet, two no-exposures-to-the-Internet control groups were included. Both control groups completed the essays without searching the Internet but one control group completed the essays immediately while the other control group completed the essays after a half-hour time period given to plan their essays. Because the control groups did not have any exposure to the Internet, they served as a measure of students' base domain knowledge prior to searching the Internet and also as a control for the half-hour time period given to the Internet search group for searching the Internet and therefore, planning their essay.

3.1. Method

3.1.1. Participants and design

In total, 150 undergraduate students participated (42 males and 108 females) in this experimental study. Of these students, 75 were senior SC students and 75 were senior ES students (M age = 21 years and 2 months, SD = 2 years and 7 months), with equal proportions of males and females drawn from each discipline. SC students were required to have taken a first-year Biology undergraduate course and to be majoring in Honours Science, Biology, Pre-optometry, or Optometry. Students in ES were required to have taken a first-year Planning undergraduate course and to be majoring in Planning, Geography, or Environmental Resource Studies. Importantly, students in SC did not differ significantly from those in ES with respect to computer experience (measured as number of hours per week spent using computers; 18.56 and 21.48 h for SC and ES students, respectively), $F(1, 149) = 1.67$, $p > .05$.

The 75 students in each discipline were assigned randomly to one of four conditions; Internet training ($n = 20$), no Internet training ($n = 20$), no-exposure-to-Internet control group with no planning time ($n = 20$), no-exposure-to-Internet control group with planning time ($n = 15$) (see Fig. 1 for a summary of the design), with equal proportions of males and females assigned to each condition. Thus, within each condition, half the participants were comprised of SC students and the other half of the participants were comprised of ES students. Students in the Internet exposure conditions (trained and untrained) were asked to search the Internet for two topics. One topic represented an area for which they had high domain knowledge (i.e., biology topic for SC students and urban environment topic for ES students) and the other topic represented an area for which they had low domain knowledge (i.e., urban environment topic for SC students and biology topic for ES students). Students in every condition completed an essay on each of the two topics. Order of topics was counterbalanced across participants, such that half the participants completed the essay corresponding to their high knowledge domain first.

3.1.2. Procedure

Each student was involved in one testing session. Students in the Internet training conditions (one for each discipline; SC and ES) started their session with the Internet self-teaching online modules that provided instruction on both basic and advanced Internet skills (adapted from the http://www.LivingInternet.com website by Stewart (2002)). Participants were asked to read the training module provided from beginning to end, and were given the option of taking notes. The program was altered so that students could not click on links between related information in order to maintain consistency in the actual information with which students were presented. This format paralleled training on Internet searching that would be experienced outside of the laboratory. After training, students proceeded to the first Internet search on the assigned topic. Instructions for the biology topic were, "For the next half-hour I would like you to search the Internet to learn about how human metabolism works. What is the process?" Those assigned to complete the urban environment topic first were told, "For the next half-hour I would like you to search the Internet to learn about major urban environmental issues that we are currently dealing with, and measures that have been proposed to deal with them."

For all participants, the computer at the start of the study session showed the Microsoft network home page (http://www.msn.com). All participants were told that they were free to search for information in any manner they chose. After completing the first search, participants were asked to complete an essay on the computer. Specific instructions for the biology topic were, "For the next 20 min I would like you to write an essay about how human metabolism works. Describe the process". Specific instructions for the urban environment topic

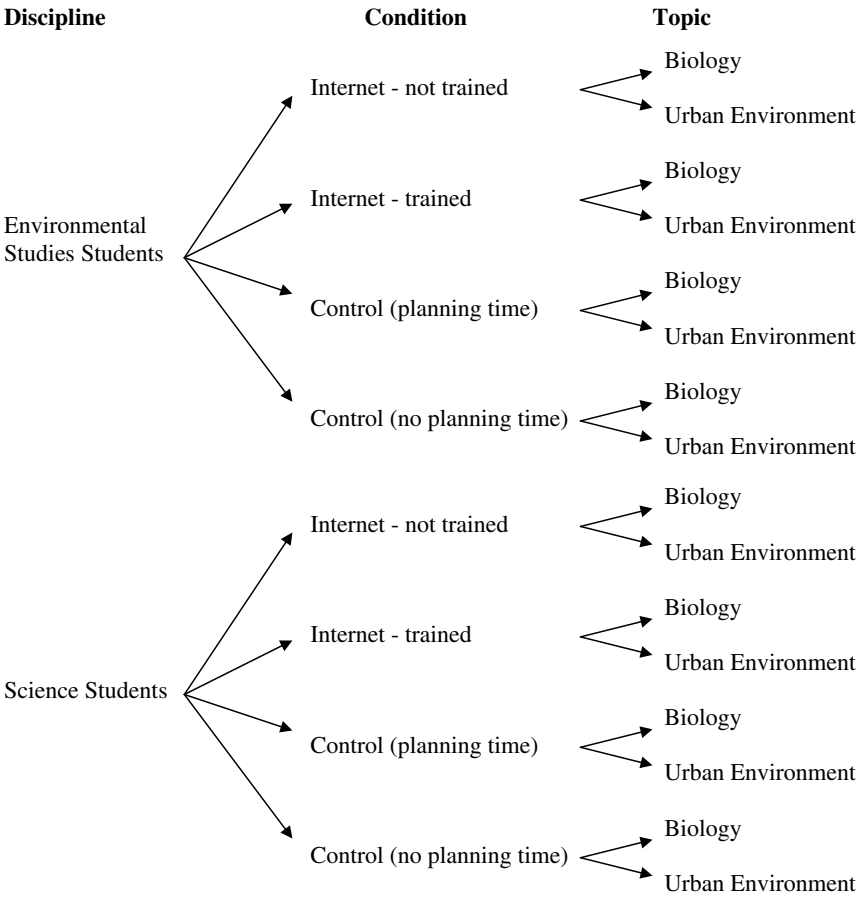

Fig. 1. Experiment 1 design.

were, "For the next 20 min I would like you to write an essay discussing major urban environmental issues that we are currently dealing with, and the measures that have been proposed to deal with them".

Students were given a 10-min break, after which they began their second Internet search followed by completion of the essay for that topic (instructions were identical to those above). Students in the untrained conditions followed all of the above instructions with the exception that they did not receive any training on how to use the Internet. Students in the no-Internet-exposure control conditions were asked to write the two essays (order counterbalanced), one for each of the two topics (i.e., human metabolism and urban environment issues). These students did not have exposure to the Internet or to any text materials related to the topic areas. One control group completed the essays immediately while the other control group completed the essays after a half-hour time period given to plan the essay.

3.1.3. Measures

A score for each essay was calculated by counting the number of acceptable factually correct statements or phrases that the participant produced that directly addressed the assigned question (see Mayer and Moreno (2002)). Twenty-five percent of the essays were scored by two raters who were highly knowledgeable with the topics but blind to the participants' level of domain knowledge, to determine inter-rater reliability. There was 82% agreement in scoring of essays, and disagreements were resolved by discussion. The remaining essays were scored by one rater.

3.2. Results and discussion

3.2.1. Impact of self-training for the acquisition of Internet searching skills

To determine whether online training in Internet search strategies enhanced the information participants retrieved from their Internet searches and then used in their essays, the first analysis compared essay performance between the Internet training and the no-Internet training conditions. Essay scores were analyzed by a repeated measures ANOVA with discipline (SC and ES) and condition as the between-subjects variables and topic as the within-subjects variable (urban environment and biology). To avoid repetition, only the comparison between conditions will be examined here; the remaining outcomes are considered below. There was no significant difference in essay scores between the Internet training, $M(SD) = 12.45\ (5.34)$, and no Internet training conditions, $M(SD) = 12.75\ (4.95)$, $F(1, 76) = .07$, $p > .05$, $\eta^2 = .00$, and no significant interactions of condition with discipline or topic. As Internet training was not a significant factor in this study, all further analyses involved the aggregated essay scores of the Internet-training and no-Internet-training conditions.

3.2.2. The role of domain knowledge

Prior to examining the role of domain knowledge on essay performance when searching the Internet compared to not searching the Internet, we first examined whether planning time differentially impacted the essay performance of the no-exposure to the Internet control

Table 1
Mean essay scores as a function of discipline, condition, and topic in Experiment 1.

Discipline Condition	Urban environment topic Mean (SD)	Biology topic Mean (SD)
Environmental studies students		
No-exposure-to-Internet-control	11.66 (5.57)	4.91 (2.98)
Internet exposure	18.30 (6.41)	6.50 (4.66)
Science students		
No-exposure-to-Internet-control	9.29 (3.54)	9.60 (6.30)
Internet exposure	11.48 (6.46)	14.13 (7.34)

groups, particularly as a function of discipline and topic. Essay scores for the two control groups were analyzed by a repeated measures ANOVA with discipline (SC and ES) and condition (planning time control group and no planning time control group) as the between-subjects variables and topic as the within-subjects variable (urban environment and biology). To avoid repetition, only the comparison between conditions will be examined here; the remaining outcomes are considered below. There was a significant main effect for condition, $F(1, 66) = 5.20$, $p < .05$, $\eta^2 = .07$, with the no planning time control group, $M(SD) = 9.76$ (4.62), outperforming the planning time control group, $M(SD) = 7.67$ (2.26). Most importantly, however, there were no significant interactions of condition with discipline or topic. Therefore, all further analyses involved the aggregated essay scores of the no-exposure-to-the-Internet conditions.

Essay scores were then compared between the Internet-exposure condition (combined trained and untrained groups) and the no-exposure-to-Internet control condition (combined no planning and planning groups). Scores were analyzed by a repeated measures ANOVA with discipline (SC and ES) and condition as the between-subjects variables and topic as the within-subjects variable (urban environment and biology).[1] There were significant main effects for condition, $F(1, 146) = 24.55$, $p < .05$, $\eta^2 = .14$, and topic, $F(1, 146) = 53.00$, $p < .05$, $\eta^2 = .27$, as well as a significant topic by discipline interaction, $F(1, 146) = 101.02$, $p < .05$, $\eta^2 = .41$. Both the main effects for topic and condition as well as the topic by discipline interaction were qualified, however, by a significant three-way interaction for topic × discipline × condition, $F(1, 146) = 11.94$, $p < .05$, $\eta^2 = .08$. See Table 1 for means and standard deviations.

Post-hoc analyses revealed that ES students who conducted Internet searches received significantly higher scores on the urban environment essays than the no-exposure-to-Internet control students, $F(1, 73) = 22.64$, $p < .05$, $\eta^2 = .24$. In contrast ES students who conducted Internet searches for the biology essay did not differ from their peers who did not use the Internet (no-exposure-to-Internet control), $F(1, 73) = 2.98$, $p > .05$, $\eta^2 = .04$. The results were reversed for the SC students. SC students who conducted Internet searches receive significantly higher scores on the biology essays than the no-exposure-to-Internet control students, $F(1, 73) = 8.08$, $p < .05$, $\eta^2 = .10$, while SC students conducting Internet searches for the urban environment essay did not differ significantly in their urban environment essay scores from the no-exposure-to-Internet control students, $F(1, 73) = 3.18$, $p > .05$, $\eta^2 = .04$. Similar to more traditional information sources, domain knowledge is an important factor in the successful retrieval and use of information from the Internet in an essay-writing task, particularly when time to search for information is more limited.

4. Experiment 2 – the relation of search behaviors to essay performance

Whereas in Experiment 1 the purpose was to assess whether domain knowledge affected essay performance after searching the Internet, in Experiment 2 the objective was to see how Internet searches might differ across knowledge areas. Specifically, this experiment examined how search behaviors were related to essay performance, that is, how the search behaviors may have contributed or detracted from essay performance. Given the limited research focused on this issue, an examination of a full battery of search behaviors was conducted in Experiment 2.

4.1. Method

4.1.1. Participants, materials, and procedure

Ten senior SC undergraduates and ten senior ES students, drawn from the same university and disciplines of students who participated in Experiment 1, participated in Experiment 2 (M age = 21.5, SD = 1.0; 6 males, 14 females). Similar to Experiment 1, students were asked to complete two Internet searches, one on human metabolism and another on urban environment issues. After completing each Internet search, students were asked to write a short essay on the topic. The procedures and materials regarding the Internet search and essay-writing task used in Experiment 1 were replicated in this study. Essays were scored for the amount of correct information that answered the question. Twenty percent of the essays were scored by two raters to determine inter-rater reliability. There was 94% agreement in scoring of essays, and disagreements were resolved by discussion. The remaining essays were scored by one rater. After completing both searches participants were asked to complete a short open-ended survey asking them to identify which topic they found easier to search (What topic did you find easier to search, and why), and which essay they found easier to write (Which essay did you find easier to write and why).

Tracking of student searches was completed using the software package WinGate 6.0. WinGate is an unobtrusive software package that downloads data on Internet use from one or more terminals. Each Internet search was coded for the number of times search engines were accessed for each topic (e.g., MSN, Google) and the number of search terms used. In addition, links that were accessed directly from the search engine results page were coded as "first links". Links that were accessed from the first links sites were coded as "second links" and links that were accessed from second links sites and beyond were called "third and subsequent links". The number of relevant/good, fair, and irrelevant/poor sites also was coded, as well as time spent on irrelevant sites (coded as time spent on irrelevant sites divided by

[1] The analysis was conducted first with gender included as a factor. Gender was not significant (largest $F(1, 142) = 3.25$, $p > .05$ for the topic × gender × condition interaction). Therefore, the analysis was rerun excluding gender as a factor.

Table 2
Search behaviors as a function of discipline and topic in Experiment 2.

Discipline Search behavior	Urban environment topic Mean (SD)	Biology topic Mean (SD)
Environmental studies students		
Number of times search engines accessed	4.40 (3.27)	4.50 (2.55)
Number of terms used in search	5.00 (3.53)	10.40 (3.69)
Number of first links	17.00 (8.47)	30.00 (6.60)
Number of second links	21.50 (6.57)	29.50 (7.40)
Number of third and subsequent links	19.80 (14.45)	10.70 (6.41)
Number of relevant/good sites	8.80 (9.41)	2.10 (1.97)
Number of fair sites	30.10 (11.93)	30.60 (8.02)
Number of irrelevant sites	0.40 (0.70)	4.40 (3.34)
Time spend on irrelevant sites	.007 (.014)	.042 (.033)
Science students		
Number of times search engines accessed	2.50 (1.96)	2.80 (1.03)
Number of terms used in search	4.80 (5.45)	7.40 (4.16)
Number of first links	10.30 (8.00)	20.00 (7.20)
Number of second links	12.10 (8.21)	22.90 (9.55)
Number of third and subsequent links	37.50 (37.59)	27.70 (27.26)
Number of relevant/good sites	9.70 (7.77)	6.50 (6.40)
Number of fair sites	34.20 (36.49)	33.90 (27.63)
Number of irrelevant sites	0.80 (1.87)	2.90 (4.79)
Time spend on irrelevant sites	.009 (.022)	.015 (.023)

Note: Number of first links = number of links that were accessed directly from the search engine results page; number of second links = number of links that were accessed from first links sites; number of third and subsequent links = number of links that were accessed from second links sites; and time spent on irrelevant sites = time spent/total time.

total time). Relevant/good sites were sites that contained information that was directly related to the topic and contained accurate, well-founded information with documented resources. Fair sites contained accurate information but less depth and sources were not documented. Irrelevant/poor sites contained information that may be inaccurate and had no foundation, and/or was completely unrelated to the topic. Twenty percent of the searches were evaluated by two raters, with 87% agreement. Disagreements were resolved by discussion and the remaining searches were evaluated by one rater.

4.2. Results and discussion

4.2.1. Essay data

Scores were analyzed using a repeated measures ANOVA with discipline (SC and ES) as the between-subjects variable and topic as the within-subjects variable (urban environment and biology). There was a significant topic $\times$ discipline interaction, $F(1, 18) = 39.06$, $p < .05$, $\eta^2 = .68$. Post-hoc analyses revealed that ES students received significantly higher scores on the urban environment essays, $M(SD) = 14.00\ (3.92)$, than on the biology essays, $M(SD) = 7.60\ (1.90)$, $F(1, 9) = 29.17$, $p < .05$, $\eta^2 = .76$. In contrast, SC students received significantly higher scores on the biology essays, $M(SD) = 15.40\ (3.31)$, than on the urban environment essays, $M(SD) = 10.00\ (3.59)$, $F(1, 9) = 13.50$, $p < .05$, $\eta^2 = .60$. Consistent with Experiment 1, high domain knowledge yielded higher quality essays.

4.2.2. Quality of Internet searches

See Table 2 for a summary of means and standard deviations. Because of nonsymmetric difference scores across topic, planned comparisons between topic areas (biology and urban environment) were conducted for each of the SC and ES disciplines using the Wilcoxon Signed-Ranks Test. For both ES students and SC students, more "first links," "second links," and irrelevant sites were accessed for the biology topic than for the urban environment topic (smallest $z = -2.05$, $p < .05$ for the "first links" comparison among biology students). In addition, ES students used more search terms for the biology topic than for the urban environment topic ($z = -2.57$, $p < .05$) and they spent more time on irrelevant biology sites than urban environment sites ($z = 2.40$, $p < .05$). These latter two results would be expected to be related; that is, increases in the use of search terms should also result in access to more irrelevant sites. Access to these sites would then result in more time spent on irrelevant sites. Overall, however, the proportion of time spent on irrelevant sites was very small. Most participants were able to access fair sites and to a lesser extent, relevant/good sites. In fact, search behaviors tended to vary by the topic of the essay rather than by domain knowledge. In other words, biology searches were different from urban environment searches, independent of whether the individuals were knowledgeable in SC or ES.

We also examined whether each specific search behavior was related to essay scores. Correlations were conducted as a function of domain knowledge and topic of essay (see Table 3 for the correlation matrix). Only one search behavior was significantly related to the quality of the essay generated. For ES students, better biology essays were associated with access to fewer irrelevant biology sites. A multivariate regression analysis examining which search behaviors best predicted essay performance, in the context of all the other search behaviors, also was conducted for each topic. For both topics, none of the search behaviors significantly predicted essay performance, although there was a trend favoring higher quality biology essays for ES students who accessed fewer irrelevant biology sites, $t(9) = -1.95$, $p = .083$.

As depicted in Table 3, the correlations of search behaviors with other search behaviors yielded no consistent patterns. One critical comparison was to examine the relation of the number of relevant/good sites accessed both within the area of expertise and outside of the area of expertise. For SC students, there was only one significant correlation, the number of search terms for the biology essay was positively correlated with the number of relevant/good sites accessed. For the ES students, the greater the number of relevant sites that were accessed for the biology area search, the greater the number of irrelevant sites that were accessed. In addition, when conducting the search in their

Table 3
Correlations among search behaviors and essay scores by topic and discipline in Experiment 2.

	Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	Biology essay score	–	.33	−.35	.01	.06	−.42	.05	.13	−.04	−.67	−.52	−.47	.07	.12	−.13	−.55	−.35	−.40	.13	.12
2	Urban environment essay score	.10	–	−.37	−.16	.28	.21	−.16	−.04	.24	−.23	−.34	.03	.34	.32	.31	.14	.18	.16	.53	.47
3	Biol # searches	−.46	.48	–	.43	−.17	.09	.33	−.48	.07	.37	−.05	.76	−.10	−.43	−.17	.21	.15	−.06	−.25	−.24
4	Biol # terms	−.24	−.07	.54	–	.22	−.10	.22	−.01	−.10	.04	−.40	.45	.37	−.21	−.21	−.02	−.15	−.08	−.16	−.04
5	Biol first links	−.30	−.43	.19	.65	–	.63	−.59	−.06	.12	.06	−.01	−.09	−.07	−.24	.06	.26	.02	.33	−.48	−.50
6	Biol second links	−.16	−.44	.17	.54	.90	–	−.24	−.44	.66	.59	.47	.01	−.14	−.30	.10	.78	.43	.68	−.28	−.31
7	Biol third and subsequent links	.06	.30	.53	.52	.18	.33	–	−.30	.45	.14	.23	.18	.50	.19	.03	.11	−.03	.15	.38	.56
8	Biol # relevant/good	−.20	.59	.66	−.05	−.01	.18	.53	–	−.43	−.75	−.10	−.21	−.16	−.08	−.19	−.43	−.55	−.19	−.03	−.10
9	Biol # fair	.01	.14	.54	.68	.36	.45	.95	.36	–	.34	.51	−.15	.13	−.15	.01	.64	.25	.64	.11	.15
10	Biol # irrelevant	.16	−.59	−.66	−.44	.08	.28	−.19	−.23	−.15	–	.39	.22	.11	.04	.19	.64	.62	.39	−.08	.02
11	Biol time irrelevant	.06	−.60	−.70	−.43	.07	.22	−.25	−.31	−.22	.97	–	−.16	−.14	−.12	.07	.47	.01	.61	−.18	−.16
12	Urban # searches	−.18	.46	.50	.30	−.14	−.07	.84	.47	.76	−.22	−.22	–	.18	−.16	.07	.19	.09	−.01	.07	.05
13	Urban # terms	.27	−.45	−.46	−.33	−.26	−.00	−.11	−.30	−.07	.77	.75	−.10	–	.75	.66	.31	.35	.40	.32	.78
14	Urban first links	.49	−.31	−.77	−.40	−.21	−.08	−.39	−.54	−.33	.73	.71	−.34	.72	–	.82	.12	.47	.22	.32	.76
15	Urban second links	−.07	−.03	−.39	.13	.22	.10	−.13	−.39	−.05	.29	.43	.02	.11	.50	–	.52	.69	.59	−.07	.47
16	Urban third and subsequent links	.16	−.01	.31	.70	.28	.35	.79	.06	.83	−.18	−.15	.55	.09	−.21	.11	–	.74	.90	−.23	.09
17	Urban # relevant/good	−.49	.08	.21	.31	−.07	−.06	.34	.04	.31	−.07	.10	.59	.16	−.15	.41	.50	–	.52	−.12	.28
18	Urban # fair	.28	.09	.18	.69	.37	.44	.76	.04	.80	−.11	−.10	.50	−.01	−.06	.32	.93	.38	–	−.26	.12
19	Urban # irrelevant	.32	−.48	−.66	−.59	−.16	.07	−.30	−.24	−.27	.94	.86	−.30	.82	.81	.12	−.32	−.26	−.27	–	.77
20	Urban time irrelevant	.38	−.38	−.68	−.62	−.20	.01	−.32	−.24	−.30	.89	.80	−.29	.73	.85	.16	−.39	−.34	−.29	.98	–

Note: Above the diagonal = environmental studies students, below the diagonal = science students; biol = biology, urban = urban environment; *ps* < .05 for *r*s > /.64.

area of expertise, ES students' access to relevant sites was correlated with increasing exploration of second, third and subsequent links. That is, there appeared to be greater branching across sites in these searches.

Importantly, there were large individual differences evident in the search behaviors. These individual differences appeared within area of domain expertise and between areas of expertise. For example, after collapsing the number of sites accessed across first links, second links, and third and subsequent links, one SC student accessed 131 sites during their biology search and 111 sites during their urban environment search. In contrast, another SC student accessed only nine sites during their biology search and 36 sites during their urban environment search. Similar variability was evident among ES students. Participants also were asked to indicate which search and which essay (i.e., biology or urban environment topic) was easier to write. All of the ES students and half of the SC students felt that the urban environment topic was easier to search than the biology topic. In addition, 8 of the10 ES students and half of the SC students felt that the urban environment essay was easier to write than the biology essay. Among the ES students then, there was a clear preference and perceived sense of ease searching and writing in the area where they had high domain knowledge. Among the SC students, half of the students identified their area of expertise as the most comfortable venue to search and write an essay.

Differences between perceived ease of search/essay-writing and actual essay scores were examined descriptively (repeated measures ANOVAs were not conducted due to the low n [$n = 5$] for the "biology topic as easiest" group). Respondents who perceived a search to be easier demonstrated higher mean scores in the essay for that area (e.g., those who perceived the biology search to be easier than the urban environment search had higher means scores for the biology essay than for the urban environment essay, M(SD) = 16.20 (2.28) and 10.00 (4.74), respectively). Similarly, respondents who perceived one essay to be easier than the other did better on the essay perceived as "easier" (e.g., when the biology essay was judged easier than the urban environment essay to write, performance on the biology essay was descriptively higher than on the urban environment essay, M(SD) = 15.00 (5.42) and 10.14 (3.02), respectively).

5. Discussion

An expected but yet striking finding in the present study was the consistent result that providing participants with 30 min to search for information on the Internet for a low domain knowledge topic did not result in better essays in comparison to their peers who had no time to search the Internet for information. Access to the Internet, therefore, may be particularly valuable as a source for information when individuals have high domain knowledge to facilitate their search (see also Park and Black (2007)). Consistent with the vast history of literature examining the impact of high domain knowledge, or expertise in other areas (e.g., Chi, 1978; Chiesi et al., 1979; Schneider et al., 1990; Voss et al., 1980), we know that expertise results in faster and more elaborate processing, limited need for external strategies, and higher retention for new but semantically related information. In the present study, even though the Internet-exposed participants writing in a low domain knowledge area had the opportunity to study material on the Internet for a full 30 min prior to writing their essays, their perfor-

mance did not differ from individuals in the control groups who had no opportunity to study at all. This finding indicates the challenges faced by less knowledgeable learners when trying to find and use novel information in the Internet when they have limited time. Past research examining the impact of domain knowledge on traditional text-searching skills also suggested that less knowledgeable individuals acquire less information from text material, even when the desired text material is readily available and accessible (see Symons and Pressley (1993)).

In light of the controversy that exposure to computer use sometimes elicits (Healy, 1998), some might argue that the findings in the present two studies suggest that exposing novice learners to the Internet as a source for information when there is limited time may provide an ineffective learning environment. On the contrary, similar to learning from other sources, retrieving information from the Internet and then using that information in a learning context may need to be supported in novices. For example, novices may need to be provided with extra time for searching the Internet (see Desjarlais and Willoughby (2007) for a study supporting that suggestion). In addition, learners may need additional scaffolding such as strategic interventions to facilitate encoding and retrieval of the novel information (Kuhn, Amsel, & O'Loughlin, 1988; Willoughby, Porter, Belsito, & Yearsley, 1999). In general, acquisition of information is a complex process and retrieving and using information from the Internet may require all of the complex supports that are associated with other sources of information such as texts, lectures, etc. (Wood, Woloshyn, & Willoughby, 1995).

Explicit training in how to conduct Internet searches is one support that is often recommended (Gergets & Hellenthal-Schorr, 2008). In the present study, providing students with access to self-teaching online training in how to search for information on the Internet did not enhance performance outcomes. This result was possibly attributable to the level of experience in the present sample of participants, who often use the Internet as part of their school work and who have had many years of exposure and opportunities for informal training in the use of the Internet. In addition, Internet search engines have become increasingly more sophisticated, and therefore only a rudimentary level of Internet search knowledge may be required to access relevant information. Indeed, the participants in the present study were able to access relevant authoritative sites fairly accurately, even when their domain knowledge was low. Some structural supports provided by search engines then, may provide a support for less experienced Internet users and individuals with low domain knowledge; however, explicit training in the specialized skills needed for sophisticated searches may be necessary to ensure high quality searches for challenging cognitive tasks (Gergets & Hellenthal-Schorr, 2008).

A particularly important outcome in the present study was the extensive individual differences found in search behaviors. Even within individuals, search behaviors could vary dramatically from one search to another and did not differ depending on whether individuals were searching for information in a high or low domain knowledge area. This finding differs from the results reported by Hembrooke et al. (2005) who found differences in search term selection between novices and experts, but these authors focused only on search term selection rather than a full battery of search behaviors and they did not assess the Internet sites that were retrieved by the participants.

At the same time, however, although domain specific knowledge did not appear to impact on search behaviors in the present study, participants' perceptions of their searches were often tied to perceived domain competence. For example, when asked to explain why one search was nominated as easier than another, participants indicated that they felt they could evaluate the quality of one set of search sites more effectively than the sites accessed in the more challenging topic, or they indicated greater knowledge in one domain than another. In some cases participants identified areas outside of their domain knowledge as easier. In some of these cases, the rationale was that there was insufficient information accessed for their domain area—again a domain knowledge issue. Perhaps domain knowledge has a more subtle, subjective impact on search behaviors that may require more extensive examination. In addition, an assessment of how search behaviors and strategies develop when using the Internet over time is needed.

6. Conclusions

The purpose of the study was to examine the role of domain knowledge when retrieving and using information from the Internet as a resource for essay tasks, as well as to investigate the quality of Internet searches and its relation to essay performance. Searching the Internet for information for 30 min enhanced essay performance relative to a control group only for the topic for which participants had high domain knowledge. In addition, although Internet search behaviors were not associated with performance, participants highlighted the importance of domain knowledge in making their searches easier. These findings suggest that domain knowledge is one factor that educators will need to pay attention to, particularly when study time is limited, in order to maximize the ability of students to successfully retrieve and use information from the Internet. A critical finding also was the large individual differences in search behaviors. Extending this research through a detailed analysis of Internet search behaviors with a large sample is needed.

References

Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading. In P. D. Pearson (Ed.), *Handbook of reading research*. NY: Longman.

Bilal, D. (2001). Children's use of the Yahooligans! web search engine:II. Cognitive and physical behaviors on research tasks. *Journal of the American Society for Information Science and Technology, 52*(2), 118–136.

Chi, M. (1978). Knowledge structures and memory development. In R. S. Siegler (Ed.), *Children's thinking: What develops?* (pp 73–96). Hillsdale, NJ: Erlbaum.

Chiesi, H., Spilich, G., & Voss, J. (1979). Acquisition of domain-related information in relation to high and low domain knowledge. *Journal of Verbal Learning and Verbal Behaviour, 18*, 257–274.

Desjarlais, M., & Willoughby, T. (2007). Supporting learners with low domain knowledge when using the Internet. *Journal of Educational Computing Research, 37*(1), 1–17.

Dias, P., & Sousa, A. P. (1997). Understanding navigation and disorientation in hypermedia learning environments. *Journal of Educational Multimedia and Hypermedia, 6*, 173–185.

Downing, R. E., Moore, J. L., & Brown, S. W. (2005). The effects and interaction of spatial visualization and domain expertise on information seeking. *Computers in Human Behavior, 21*, 195–209.

Foltz, P. W. (1996). Comprehension, coherence, and strategies in hypertext and linear text. In J. Rouet, A. Dillon, & R. J. Spiro (Eds.), *Hypertext and cognition* (pp. 109–136). Mahwah, NJ: Erlbaum.

Gergets, P., & Hellenthal-Schorr, T. (2008). Competent information search in the World Wide Web: Development and evaluation of a web training for pupils. *Computers in Human Behavior, 24*, 693–715.

Harris, J. (1996). Information is forever in formation; knowledge is in the knower: Global connectivity in K-12 classrooms. *Computers in the Schools, 12*, 11–22.

Healy, J. M. (1998). *Failure to connect: How computers affect our children's minds – for better and worse*. NY: Simon & Shuster.

Hembrooke, H. A., Granka, L. A., & Gay, G. K. (2005). The effects of expertise and feedback on search term selection and subsequent learning. *Journal of the American Society for Information Science and Technology, 56*(8), 861–871.

Hess, B. (1999). Graduate student cognition during information retrieval using the World Wide Web: A pilot study. *Computers and Education, 33*, 1–13.
Jacobson, M. J., & Spiro, R. J. (1995). Hypertext learning environments, cognitive flexibility, and the transfer of complex knowledge: An empirical investigation. *Journal of Educational Computing Research, 12*, 301–333.
Kuhn, D., Amsel, E., & O'Loughlin, M. (1988). *The development of scientific thinking skills*. San Diego: Academic Press.
Kuiper, E., Volman, M., & Terwel, J. (2008). Integrating critical web skills and content knowledge: Development and evaluation of a 5th grade educational program. *Computers and Human Behavior, 24*, 666–692.
Lidstone, J., & Lucas, K. B. (1998). Teaching and learning research methodology from interactive multimedia programs: Postgraduate students' engagement with an innovative program. *Journal of Educational Multimedia and Hypermedia, 7*, 237–261.
Marchionini, G. (1995). *Information-seeking in electronic environments*. New York: Cambridge University.
Mitchell, T. J. F., Chen, S. Y., & Macredie, R. D. (2005). Hypermedia learning and prior knowledge: Domain expertise vs. system expertise. *Journal of Computer Assisted Learning, 21*(1), 53–64.
Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology, 93*(1), 187–198.
Mayer, R. E., & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction, 12*, 107–119.
Park, Y., & Black, J. B. (2007). Identifying the impact of domain knowledge and cognitive style on web-based information search behavior. *Journal of Educational Computing Research, 38*(1), 15–37.
Perry, N., VandeKamp, K., Mercer, L., & Norby, C. (2002). Investigating teacher–student interactions that foster self-regulated learning. *Educational Psychologist, 37*(1), 5–15.
Pintrich, P. (Ed.). (1995). Current issues in research on self-regulated learning: A discussion with commentaries. *Educational Psychologist, 30*(4), 171–232 [Special issue].
Pritchard, A., & Cartwright, V. (2004). Transforming that they read: Helping eleven-year-olds engage with Internet information. *Literacy, 38*(1), 26–31.
Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. Spiro, B. Bruce, & W. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 139–175). Hillsdale, NJ: Erlbaum.
Schneider, W., Korkel, J., & Weinert, F. (1990). Expert knowledge, general abilities, and text processing. In W. Schneider & F. Weinert (Eds.), *Interactions among aptitudes, strategies, and knowledge in cognitive performance* (pp. 235–250). New York: Springer-Verlag.
Sewell, D. F. (1990). *New tools for new minds: A cognitive perspective on the use of computers with young children*. New York: St. Martin's Press.
Slatin, J. M. (1990). Reading hypertext: Order and coherence in a new medium. *College English, 52*, 870–883.
Stewart, W. (2002). The living internet. Accessed on April 25, 2008. Retrieved from <http://www.livinginternet.com/>.
Symons, S., MacLatchy-Gaudet, H., Stone, T. D., & Reynolds, P. L. (2001). Strategy instruction for elementary students searching informational text. *Scientific Studies of Reading, 5*, 1–33.
Symons, S., & Pressley, M. (1993). Prior knowledge affects text search success and extraction of information. *Reading Research Quarterly, 28*(3), 250–259.
Tu, Y.-W., Shih, M., & Tsai, C.-C. (2008). Eighth graders' web searching strategies and outcomes: The role of task types, web experiences and epistemological beliefs. *Computers and Education, 51*, 1142–1153.
Voss, J., Vesonder, G., & Spilich, G. (1980). Text generation and recall by high-knowledge and low-knowledge individuals. *Journal of Verbal Learning and Verbal Behaviour, 19*, 651–667.
Willoughby, T., Porter, L., Belsito, L., & Yearsley, T. (1999). Use of elaboration strategies by students in grades 2, 4, and 6. *Elementary School Journal, 99*, 221–232.
Willoughby, T., Wood, E., & Khan, M. (1994). Isolating variables that impact or detract from the effectiveness of elaboration strategies. *Journal of Educational Psychology, 86*(2), 279–289.
Willoughby, T., Wood, E., & Kraftcheck, E. R. (2003). When can a lack of structure facilitate strategic processing of information? *British Journal of Educational Psychology, 73*, 59–69.
Wood, E., Woloshyn, V. E., & Willoughby, T. (1995). *Cognitive strategy instruction for middle and high schools*. Cambridge: Brookline Books.
Zimmerman, B. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology, 81*, 329–339.