Int. J. Human-Computer Studies (2002) **57**, 1–26
doi:10.1006/ijhc.1011
Available online at http://www.idealibrary.com.on IDEAL®

The effects of graphic organizers giving cues to the structure of a hypertext document on users' navigation strategies and performance

RACHEL MICHAEL NILSSON AND RICHARD E. MAYER
Department of Psychology, University of California, Santa Barbara, CA 93106, USA. email: nilsson@psych.ucsb.edu, mayer@psych.ucsb.edu.

(Received 6 November 2001 and accepted in revised form 6 June 2002)

Two experiments examined the effects of graphical organizers on users' navigation of a 150-page hierarchical website of aquatic animals. In Experiment 1, users were given either a non-clickable map (map group) or no map (no-map group) and answered 30 questions by searching the website. The map group was more efficient (visited fewer pages) on the first 20 questions (learning phase) but the no-map group was marginally more efficient on the last 10 questions (test phase), and displayed more flexible search strategies. In Experiment 2, users were either given a simplified organizer locating the current page in the website (explicit group) or an alphabetized list of superordinate pages (implicit group). The task from Experiment 1 was repeated. No differences in efficiency were found, but the explicit group was faster than the implicit group in the test phase. The results depended on individual differences in spatial skills. These results suggest a tradeoff between organizers that are useful initially and those that promote structural learning.

KEYWORDS: graphical organizer; hypertext; navigation strategy.

1. Introduction

Hypertext documents are used in educational and business applications. These include data storage and retrieval in the place of traditional databases and corporate web environments (Vora & Helander, 1997). Unfortunately, our understanding of how humans mentally represent the information available to them and learn to navigate the hypertext systems has not kept up with the technology. Attempts to find specific information can be marked by disorientation and cognitive overload (Conklin, 1987), resulting in frustration for users of the web. Nickerson and Landauer (1997) suggest that a better understanding of how humans conceptualize information spaces would increase the usability of computer software.

In this study, we examine how information about the structure of a hypertext document affects how users represent and navigate the document. Specifically, these two experiments were conducted to determine whether the structural information

1071-5819/02/$ - see front matter

contained in graphic organizers aids users in learning the structure of the website leading to improvements in navigation.

1.1. DESCRIPTION OF THE HYPERTEXT DOCUMENT

The hypertext document used in this study was a 150-page website of fish and other aquatic animals, which the participants used to answer 30 questions about specific animals. Of the 150 pages, 67 pages (i.e. animal pages) gave information about a specific animal, and the other 83 pages (i.e. group pages) gave information about a group of fish sharing a common biological classification, diet, or habitat. The group pages were arranged into three hierarchies, with more general groups linked to more specific subordinate pages, down to the level of individual fish. These group pages provided the means for the user's navigation. Across the two experiments described in this paper, four versions of the website were presented to participants, and their navigation performance on the different versions was compared. Three versions included different organizers with different structural information and the fourth contained no organizer at all.

Two dependent measures used across the different versions of the website were the efficiency of participants' performance, calculated separately for the first 20 questions, which were considered to be the learning phase, and the last 10, which were considered to be the test phase. This allowed us to examine the changes in the performance of participants over time. Also, the search strategies of the users, their self-reports of usability, and the accuracy of their searches were examined on the last 10 questions to pinpoint differences in performance that might reflect different mental representations of the website's structure.

1.2. ROLE OF MAPS

One navigational aid often found in hypertext documents is a map of the document. A map (as used in this paper) is a graphic representation of a hypertext document, in which pages of the document are represented by visual objects (whether simply the title of a page or an icon representing a page) and the links between pages are represented by lines or arrows connecting the visual objects.

Several studies have shown that maps can improve user navigation. Monk, Walsh and Dix (1988) found that the inclusion of a map increased the number of questions answered per hour using a hypertext document. Similarly, Hammond and Allinson (1989) found that the inclusion of a map in a hypertext document made users more efficient (higher ratio of new pages visited over total pages visited) in searching the document. Simpson and McKnight (1990) found that users given an outline of a hypertext document visited fewer extraneous pages in searching the document than those given an alphabetical list of topics. Tripp and Roby (1990) found that the inclusion of either a graphical advance organizer or a spatial metaphor for the organization of a hypertext lexicon increased the amount users learned from the lexicon. While their study looks at vocabulary learning rather than navigation performance, Tripp and Roby attribute the poorer performance of the no-organizer group to students becoming disoriented and re-viewing pages they had already seen, rather than exploring more fully. McDonald and Stevenson (1998*b*) found that the inclusion of a map led to better performance on search tasks than a contents list or bare hypertext. In a meta-analysis by Chen and Rada (1996) maps were found to be consistently helpful for new users, with an effect size of $r = 0.38$, unlike indices and tables of contents, which were not helpful.

One possible explanation as to why some maps help users is that they reduce the cognitive load of navigation, allowing the user to learn the structure of the website (Tripp & Roby, 1990; McDonald & Stevenson 1998*b*). Cognitive load results from an individual's attempt to maintain and consider unrelated concepts in working memory (Sweller, 1988). Maps give information about the structure of the website that may cue important connections between concepts, helping users integrate their knowledge. Maps also provide an external representation of the website, freeing cognitive resources otherwise employed in retaining this information. According to cognitive load theory, maps lower the lower cognitive load of navigation, leading to better learning of the structure of the website. Thus, cognitive load theory predicts that the benefits of including a map would increase over time.

On the other hand, maps might be useful only in that they tell users where to find the next answer, and how to get from where they are to where the answer is. In such a case, navigation requires little thought or integration of information into a coherent representation. For example, Jonassen and Wang (1993) found that providing a graphical browser did not improve participants' performance on judging relationships between ideas, but that requiring the participants to design their own semantic network did. Similarly, Shapiro (1998) found that participants given explicit information about the structure of information in a hypertext document performed worse on an integrative essay test than those given no such information. Shapiro suggests that these results may be due to elaborative, constructive learning processes on the part of participants given no structural information. Similarly, the work of Mannes and Kintsch (1987) found that forcing participants to consider more deeply the structure of a document (by providing a prior organizer with a different organization) led to the participants gaining a deeper understanding of the materials, and Dee-Lucas and Larkin (1995) found that when students are given a specific learning goal, a hierarchical interactive overview of a hypertext document provides no benefits. If deep, elaborative processes are important for meaningful learning, then having some kinds of maps may be useful initially but may not lead to improvement in navigation over time. We refer to this as the active learning theory. Active learning theory predicts that maps will lead to initial improvements in navigational ease but not to long-term improvements in understanding the structure of the website.

To distinguish between the two possibilities described above requires that performance be examined over time, as the user has a chance to learn (or not learn) the structure of the website and to develop search strategies. In the following studies, therefore, we break up the hypertext navigation task into two phases, a learning phase (first 20 questions) and a test phase (last 10 questions). The purpose of these studies is to determine what effect structural cues contained in graphical organizers have on students' ability to learn to navigate a hypertext document.

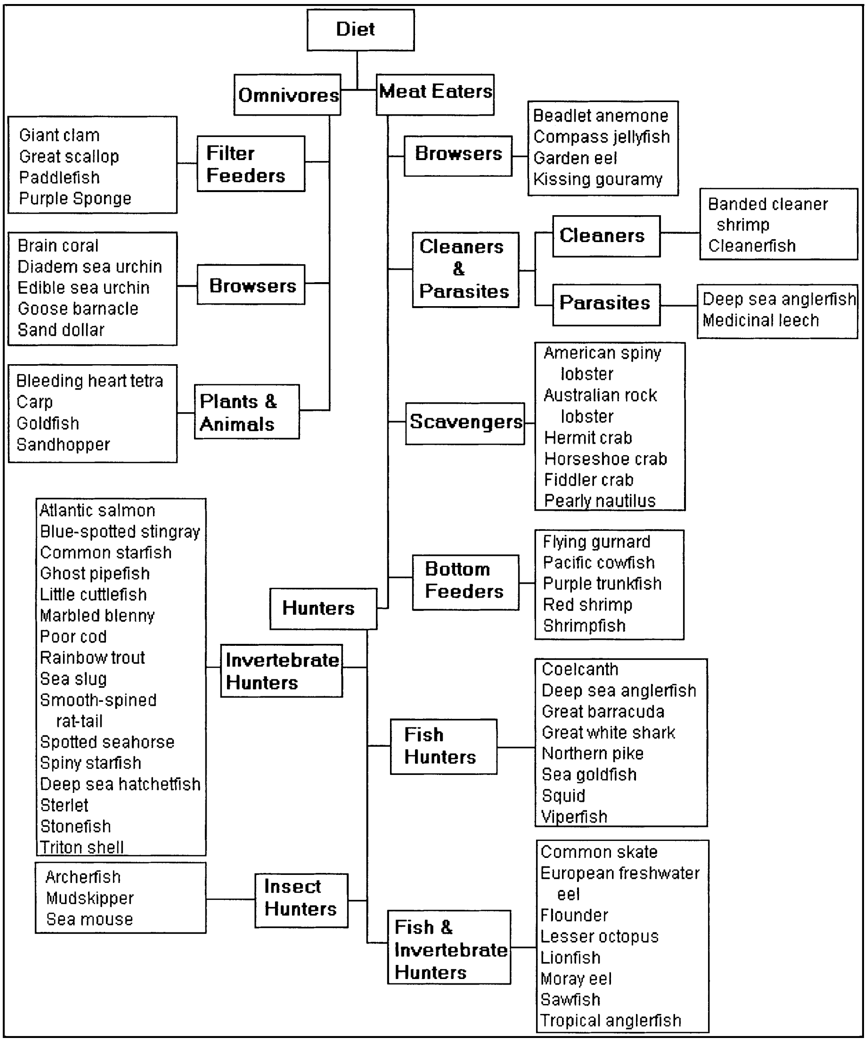

FIGURE 1. Map of diet group pages and animal pages.

2. Experiment 1

Experiment 1 was created to test the hypothesis, based on the cognitive load theory, that maps will improve the user's navigation performance over time. The alternate hypothesis, based on the active learning theory, is that maps simply tell users where to go, and will not help users improve over time.

Two versions of the website were created. One group of users (Map Group) performed the task using a version of the website containing three maps of the website, one of which can be seen in Figure 1. Three maps were used since the size and interconnectedness of the website made a single map visually confusing and unwieldy. These maps were accessible from the animal pages via a link on the page, and popped up to replace the animal page. The user could then return to the animal page to continue the navigation task. The other group (No-Map Group) used a version of the website that was identical except that no maps were included.

The maps used were non-navigable; that is, once the participant viewed the map, he or she needed to return to the animal page to navigate via the hierarchical link structure. This was done for several reasons. The first is that the past research on maps and hypertext navigation performance, such as by Simpson and McKnight (1990) and Monk, *et al.* (1988), used non-navigable maps. The second reason is that we were interested in the effects of providing the users with information about the structure of the website. Providing one group, but not the other, with a different means of navigation in addition to structural information would have been a confound.

The cognitive load theory predicted that the map group would take less time and visit fewer pages to answer questions than the no-map group at the beginning of the task, and that the gap in performance between the groups would widen as the two groups got more experience with the website. It also predicted that the map group would back up (return to the most recently viewed page—a measure of error) less often after they have had a chance to learn their way around the website. In addition, the cognitive load theory predicted that the no-map group, being overloaded, would attempt to create a simplified mental representation of the document that only includes one or two of the three hierarchies, and would only search one or two of them during the test phase, ignoring the third.

The active learning theory, on the other hand, predicted that while the map group may be more efficient during the learning phase, they will not show the same improvement over time as the no-map group. It also predicted that the map group would back up more often than the no-map group, and that the map group would be the one creating a simplified mental representation of the document.

2.1. METHOD

2.1.1. Participants. The participants were undergraduates at the University of California, Santa Barbara. They participated for class credit, and consent was obtained prior to participation. In total, 66 students (51 women, 15 men) participated in the experiment, although computer problems and participants' failure to follow directions led to the exclusion of 13 participants, leaving 11 men and 42 women.

The mean SAT scores for the map group ($M = 1218.41$, S.E. $= 27.96$) did not differ from that of the no-map group ($M = 1174.78$, S.E. $= 21.63$), t (43) $= -1.24$, $p = 0.22$. The mean amount of time spent on the World Wide Web did not differ for the map group ($M = 2.62$, S.E. $= 0.62$) and the no-map group ($M = 3.13$, S.E. $= 0.58$), t (52) $= 0.60$, $p = 0.55$. In addition, the ratio of males to females in the map group (6/21) was the same as that in the no-map group (6/21), χ_2 (1, $N = 54$) $= 0.00$, $p = 1.00$. The proportion of participants with web design experience did not differ significantly for the map (4/27) and no-map group (6/27), χ_2 (1, $N = 54$) $= 0.49$, $p = 0.484$. The mean age for the map group ($M = 18.26$, S.E. $= 0.14$) did not differ from that of the no-map group ($M = 18.56$, S.E. $= 0.22$), t (52) $= 1.14$, $p = 0.26$. Finally, the mean academic year for the map group ($M = 1.22$, S.E. $= 0.08$) did not differ from that of the no-map group ($M = 1.52$, S.E. $= 0.19$), t (52) $= 1.45$, $p = 0.15$.

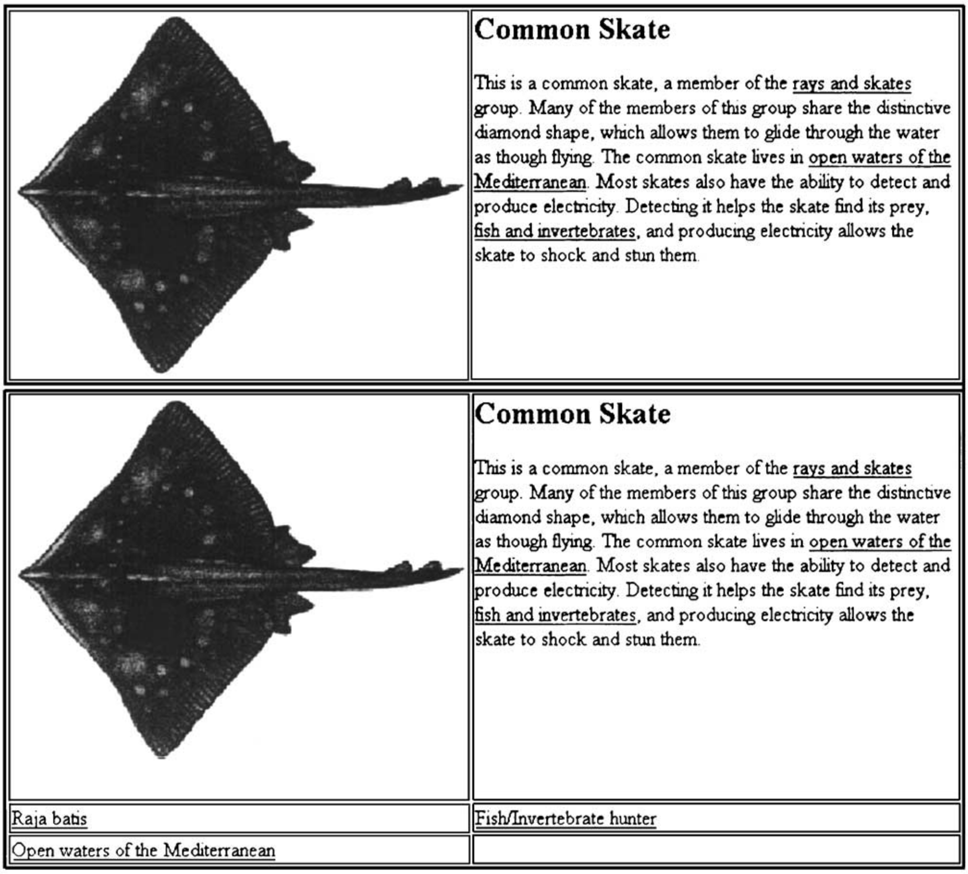

FIGURE 2. Two versions of the common skate page from Experiment 1: bottom version is from map condition; top version is from no-map version.

2.1.2. Materials and apparatus. Two versions of an aquatic animals website were created to serve as stimuli—a map version and a no-map version. Both versions contained the same number of animal and group pages, in addition to which the map version contained three map pages. Figure 2 shows an example animal page from each version of the website. The page contains information about a particular animal, including its habitat, diet, biological classification, a picture of the animal, and other interesting facts. The animal pages were the same in both versions, except that in the map version each of the animal pages contained three terms at the bottom of the frame linking to the three maps of the website.

In both versions of the website, each animal page was linked to three group pages—one for the animal's habitat, one for its diet, and one for its biological classification. Figure 3 shows one of these group pages. Each group page described the animals in the group, and linked "upwards" to a superordinate group page and "downwards" to subordinate pages, either group pages or animal pages. This structure allows the

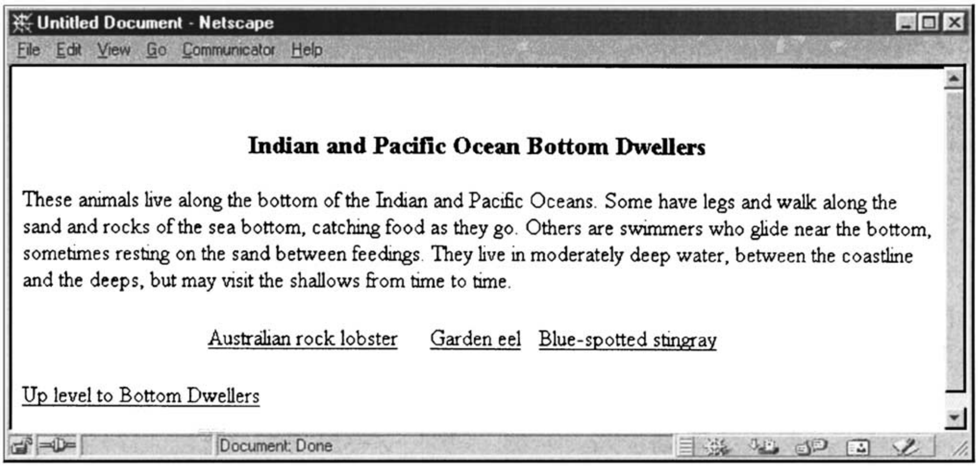

FIGURE 3. Group page, which appears the same in both versions of the website.

participant to move from one animal to another by moving up and down in each hierarchy. For example, the Giant Clam page links to the Bivalves page, which then links down to other animal pages and up to the Molluscs group page.

The diet hierarchy can be seen in its corresponding map, as shown in Figure 1. This map is one of the three used in the experiment. The maps mirror the structure of the website. A box represents each group page in the hierarchy, each line represents a link. The animal pages were too numerous to get their own boxes, so all of the animals linked to one group page were put into one box.

A list of 30 questions was created to give the users an opportunity to experience the website and to allow us to assess their knowledge. These questions were typed onto half sheets of paper, which were stapled into a booklet for each participant. The first 10 questions asked about the diets, habitats, and classifications of 10 animals. This was intended to call the participants' attention to these three attributes, and give them a chance to use all three hierarchies. In addition, following the link that explained the answer to the question led to a group page containing the name and link to the animal that was asked about in the next question. For example, participants were asked, "What is the diet of the (filter feeder) giant clam?" Participants who followed the link to filter feeders to find out more about the term then found a link to the next animal they had to look up (another filter feeder) on that page. This was intended to give the participants early success in using the system. The next 10 questions asked about 10 randomly selected animals. For example, participants were asked, "What made the breeding habits of the European freshwater eel so mysterious?" These questions were intended to give participants more experience with the website.

We designed the last 10 questions to assess the participants' conceptions of the information space and how to navigate through it. Like the middle 10 questions, these questions asked about information other than diet, habitat, or classification. Unlike the middle 10 questions, the next animal was always one jump away through one of the three hierarchies.

For example, participants were asked, ''How do ghost pipefish catch their food?'' and then ''How do smooth-spined rat-tails communicate?'' Both of these fish hunt invertebrates, and are thus one jump apart through the diet hierarchy. The next question is ''Why is it a good thing male deep-sea anglerfish are small?'' Both the smooth-spined rat-tail and the deep-sea anglerfish are deep-sea dwellers, and thus only one jump apart through the habitat hierarchy.

Two questionnaires were created for use in this study. The background questionnaire, which was given before the experiment began, asked participants to specify their age, major, academic status, gender, and SAT scores. Participants were also asked about their computer use (''How many hours a week do you spend using computers?'' ''How many hours a week do you spend using the World Wide Web?', and ''Have you ever made your own webpage?''). The questions were printed out on a single sheet of paper to be completed by the participants.

A comprehensibility questionnaire asked participants to rate the statements ''I think I know how the information on this website is organized'' and ''I could use this website to find out more information about other fish'' on a 5-point Likert scale from 1 (not at all) to 5 (very well). This questionnaire was printed out on a half sheet of paper and participants were asked to circle the number corresponding to their response.

The instructions for the experiment were also printed out on a single sheet of paper. These were designed in a tutorial-type format, and are described more fully below.

The experiment was run on two computer systems, a MacIntosh G3 and a Power Mac 7600, each with a 17 in color monitor. The website was stored on a local network, and accessed using Netscape 3.01.

2.1.3. Procedure. Upon entering the study, participants were randomly assigned to either the map or no-map condition. Participants were tested in separate rooms.

First, participants completed the background questionnaire. They were then seated in front of the computer and given printed instructions for using the website. Participants were told they had 10 min to read the instructions and explore the website freely, at which time they would be given another task to complete. The printed instructions highlighted features of the website in a tutorial-style format. The participants were asked to click on different links and told what these links signified. The purpose of the tutorial was to make sure the participants understood that the blue, underlined text served as links, to introduce the group pages, and to introduce the three hierarchies and their structure. In addition, participants in the map condition were also introduced to the three map pages and their function, and were shown how to access the maps from each animal page.

When the 10 min had expired, participants were given the packet of questions and asked to complete them by looking the answers up on the website. The participants were asked to complete the questions in order, without looking forward or backward. In addition, they were asked not to use the ''back'' button on the Netscape browser, due to technical difficulties, but rather to return to the previous page using the in-text links. The one exception to this was that the participants in the map condition were allowed to use the back button to return from the map pages to the animal page from which the map page had been accessed. While the participants were not told specifically that they should learn their way around, the length of the task motivated participants to learn the website and to develop search strategies.

The participants were also instructed to ask the experimenter any questions they had during the experiment. Participants then began the task. At 20 min intervals, the experimenter approached the participants to ask if there were any problems, and assistance was given to those participants who asked for it. Once each participant finished the task, he or she was given the comprehensibility questionnaire. The entire process took between 90 min and 2 h.

2.2. RESULTS

The access log generated by the network web server provided a list of what documents each participant's computer requested, and at what time. This provided us with the data needed to track each participant's path through the website, and was used to compute the statistics below.

The cognitive load theory predicted that the map group would be more efficient in the learning phase, and that they would improve more than the no-map group from the learning phase to the test phase. The active learning theory predicted that the map group would be more efficient in the learning phase but that the no-map group would improve more over time. This prediction was tested using a 2×2 mixed ANOVA.

2.2.1. Efficiency of performance. The efficiency measures were the mean number of pages each participant viewed per question (hits per question) and the average time per question for both the first 20 and last 10 questions. Figure 4 shows the mean number of hits per question for each group on the first 20 questions (learning phase) and the last 10 questions (test phase). An interaction between the phase of the experiment and the treatment condition was observed, $F(1,52) = 12.03$, M.S.E. = 285.92, $p < 0.01$. Simple main effects indicated that the map group ($M = 14.40$, S.E. = 1.18) was initially more efficient than the no-map group ($M = 18.49$, S.E. = 1.15), $F(1,51) = 6.16$, M.S.E. = 221.61, $p < 0.05$, but during the test phase the no-map group ($M = 9.89$, S.E. = 1.04) had a marginally significant advantage over the map group ($M = 12.37$, S.E. = 1.06), $F(1,51) = 2.77$, M.S.E. = 81.48, $p = 0.10$. Overall, the average number of hits per question was higher in the learning period ($M = 16.45$, S.E. = 0.82) than during the test period ($M = 11.13$, S.E. = 0.75), $F(1,52) = 31.55$, M.S.E. = 749.95, $p < 0.00$. Finally, the map group ($M = 13.38$, S.E. = 0.89) did not average significantly fewer hits overall than the no-map group ($M = 14.19$, S.E. = 0.88), $F(1,52) = 0.41$, M.S.E. = 17.17, $p = 0.52$. These results are inconsistent with our predictions. While the map was initially useful, it did not lead to improvements in performance over time.

Technical difficulties prevented the use of time data collected on one of the two machines, so data from 24 participants (13 in the map group, 11 in the no-map group) were analysed.

Again, the cognitive theory predicted that the map group would be faster than the no-map group in both phases, and that the gap would widen over time, while the active learning theory predicted the same initial difference but with the gap narrowing over time. Figure 5 shows the mean time (in seconds) for the groups in the learning and test phases. As with the average number of hits, a gain was found from the learning period

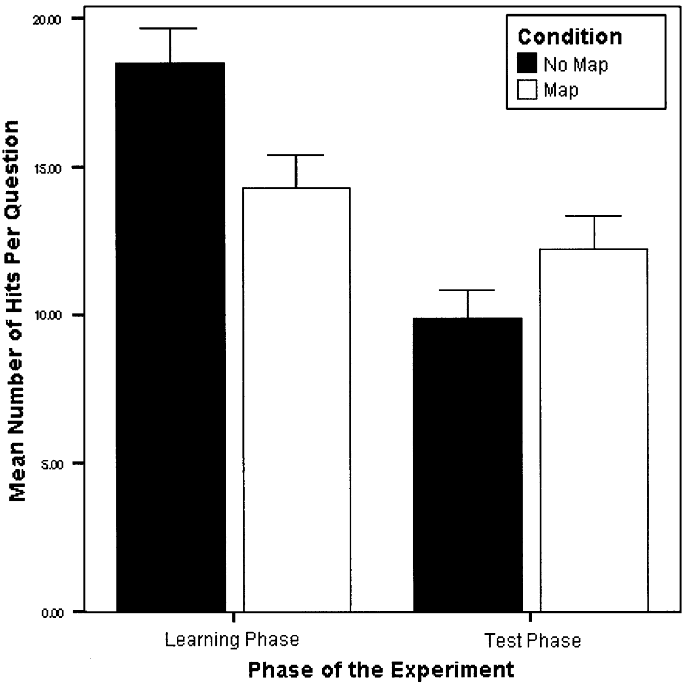

FIGURE 4. Mean number of hits per question (+S.E.) as a function of the phase of the experiment and treatment condition—Experiment 1.

(M = 107.98, S.D. = 6.32) to the test period (M = 78.55, S.D. = 5.44), $F(1,22) = 16.23$, M.S.E. = 10322.56, $p < 0.01$. No significant difference was found between the conditions, $F(1,22) = 0.20$, M.S.E. = 199.74, $p = 0.66$, and there was also no significant interaction of the two variables $F(1,22) = 0.59$, M.S.E. = 199.74, $p = 0.45$. The trend for these results appears to follow the same pattern as the hits data above, with the no-map group starting at a disadvantage but ending up faster than the map group.

2.2.2. Accuracy of performance. In addition to these overall measures, the number of backups performed during the test phase was recorded and analysed using a t-test. This was done by counting the number of times each participant, after moving to a new page, proceeded to return to the page he or she had just been looking at. We considered this to indicate an error in searching on the part of the participant. One participant was dropped from this analysis only, being an extreme outlier on this measure.

Cognitive load theory predicted that the map group would perform fewer backups than the no-map group while the active learning theory predicted the opposite. The mean number of backups for the map group (M = 18.25, S.D. = 3.58) was not significantly different from the mean number of backups for the no-map group (M = 14.40, S.D. = 2.77), $t(51) = -0.12$, $p = 0.91$. Because the scores on this measure were skewed, a Mann–Whitney U test was also run, but again, no significant differences were found, $U = 350.50$, $p = 0.99$. Neither group was more accurate in their searches.

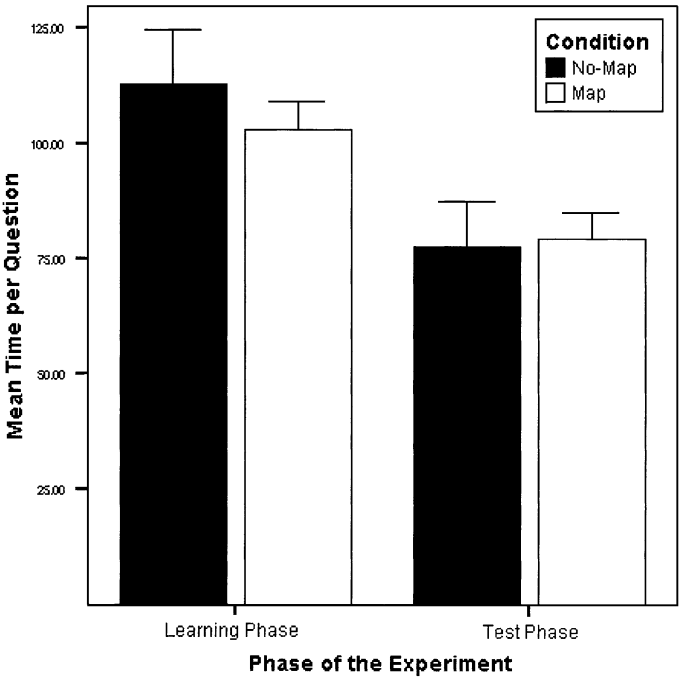

FIGURE 5. Mean time per question (+S.E.) as a function of the phase of the experiment and treatment condition—Experiment 1.

2.2.3. Quality of models. The cognitive load theory predicted that the no-map group would be overloaded and therefore limit the number of hierarchies they searched, accepting a decrease in efficiency in order to decrease the cognitive load of the task. The active learning theory predicted that the map group, being less involved in the task, would fail to consider all three hierarchies when searching.

For each participant, we calculated a preference score, which represents the degree to which that participant restricted his or her searches to only one or two hierarchies. For each question, we noted which hierarchy or hierarchies had been used by the participant searching for the animal in question. We then totaled the number of such notations, such that a hierarchy used as part of answering all 10 questions would have a total of 10. Then we subtracted the smallest of these hierarchy totals from the largest, and divided the result by the sum of all three totals, then multiplied by 100. This created a preference score that captured the degree to which a participant used only one or two hierarchies to the exclusion of other(s), while controlling for the length of search. A participant who searched one and only one hierarchy across all questions would have a score of 100. A participant who used all three hierarchies equally would have a score of zero.

Contrary to our predictions, the map group (M = 50.27, S.E. = 6.10) had a higher preference score than the no-map group (M = 32.97, S.E. = 5.23), t (51) = −2.157, $p < 0.05$. Because these scores were skewed, a Mann–Whitney U test was also run, and the same difference found, U = 232.50, $p < 0.05$. The map group was more likely to pick one hierarchy and search it repeatedly, despite the fact that this limitation would hinder the search for information. This suggests either the no-map group was less overloaded than predicted, or that they were more involved in the task, and therefore more willing to take on the cognitive load of dealing with all three hierarchies.

2.2.4. Strategy use. Once we realized that the results of the tests discussed above did not match the predictions of the cognitive load theory, which we had favored prior to the data collection, we began to look at the participants' search strategies using the server log. Three questions of the final 10 were chosen for closer analysis. These three questions all contained cues in the name of the animal that could potentially help the participant choose a comparatively short path through the website. For example, the participant is asked to navigate from the smooth-spined rat-tail to the deep-sea anglerfish. The rat-tail is obviously a deep-sea fish, with bulbous eyes and pale coloring. The text discussing the rat-tail also identifies it as living in deep ocean waters, and contains a link to "Deep Ocean Waters." Since the next animal name contains the word "Deep-Sea," this would presumably be a good link to follow.

Participants responded to these questions in one of two different ways. Some chose to use cues like the name of the animal, which to us signaled a greater involvement in the task. Following the link from a deep-sea animal to a page labeled "Deep Ocean Waters" does not require detailed knowledge of what pages are where, but rather thoughtful involvement in the task of interacting with the website. Other participants ignored these cues, generally to follow a more stereotyped search routine, such as looking up the animal on the map of the participant's favorite hierarchy, and then following these directions. This was seen as an indication the participant was less cognitively involved in the task, relying on knowing what pages are where rather than knowing how to interact with the website.

We therefore divided the participants into groups based on their navigational strategies in response to such questions. Structural strategies were (a) looking up the next animal on one of the maps and then following that path exactly, (b) following a direct path through the one hierarchy the participant seemed to have memorized, and (c) conducting an exhaustive search of one hierarchy (almost always the same hierarchy for each question). These strategies (whether efficient or not) relied on knowing the structure of the website, but did not indicate involvement in the task.

Task strategies were (a) taking the name of the next animal as a cue for performing the search, and (b) searching for the next animal by moving out one page into each hierarchy, then returning to the animal if the desired animal was not found (an effective strategy, since the last 10 animals could all be found rapidly in this fashion). These strategies did not require the participant to know what pages are where in the website, but did require attention to the task. Of course, this is not to say that these participants lacked a detailed mental representation of the website, but rather that their strategies did not rely on such a representation.

A 2×2 chi-square was run comparing strategy choices and treatment condition. A greater proportion of the no-map group (20 out of 27) predominantly chose task strategies than in the map group (9 out of 26), χ^2 $(1, N = 53) = 8.32$, $p < 0.01$. As with the preference scores reported above, these results suggest that the no-map group, rather than retreating cognitively and simplifying the task, was more involved in and willing to commit their cognitive resources to the task.

2.2.5. Self-reports of usability. Both the cognitive load theory and the active learning theory predicted that the map group would rate the website as more usable than the no-map group. The mean response to the statement "I think I know how the information on this website is organized" for the map group ($M = 4.15$, S.E. $= 0.17$) was marginally higher than for the no-map group ($M = 3.74$, S.E. $= 0.11$), $t\ (52) = -1.95$, $p = 0.056$. The mean response to the statement "I could use this website to find out more information about other fish" was no different for the map group ($M = 4.52$, S.E. $= 0.13$) than for the no-map group ($M = 4.26$, S.E. $= 0.17$), $t(52) = -1.47$, $p = 0.15$. The map group seemed to find the website marginally more comprehensible.

2.3. DISCUSSION

2.3.1. Summary of results. Overall, the results were more consistent with the active learning hypothesis. The map group was initially more efficient in navigating the website, as both theories predicted. However, the no-map group showed greater improvement from the learning phase to the test phase and was therefore marginally more efficient during the test phase. In addition, the no-map group attended to the task and showed a greater willingness to search all three hierarchies. This pattern of results is more consistent with the active learning theory.

Put together, these data create clear profiles of the two different groups. Participants given a map found the structure of the document more comprehensible. Using the map, the participants narrowed in on one or two hierarchies to focus on, and quickly developed a search strategy, either relying on the map, or trusting themselves to memorize that one hierarchy. These search strategies showed little or no improvement over time, possibly because they were adequate to the task and required no further elaboration. In short, these participants were not forced to process their task deeply, so they did not.

Participants not given a map found themselves in a different situation. While the names of the links and the pages gave a number of cues to the participants, they were forced to search for the next animal rather than simply look up the route. This led to the development of task strategies that focused on using cues in the names of the pages and a more general awareness of the types of links between pages rather than the details of the hierarchies. These participants were forced to process their task deeply, and they did so.

2.3.2. Implications for the theories. The results of the experiment support the theory that learners more actively involved in a task learn more from it. While the negative effects of graphical organizers have been shown on content post-tests in the past (Mannes & Kintsch, 1987; Jonassen & Wang, 1993; Dee-Lucas & Larkin, 1995; Shapiro, 1998), we have now found it on navigation of the document as well.

The implications for the cognitive load theory are less clear cut. The map group was initially more efficient and rated the system as being more comprehensible, which suggests that there was a reduction in cognitive overload for this group. This finding also rules out the possibility that the maps were too difficult to use. If this had been the case, there should have been no advantage for the map group during the learning phase, and no difference on the self-report measures. It simply appears that any benefits from the structural information contained in the map were overshadowed by the negative effects of the map decreasing participants' involvement in the task.

While the results of Experiment 1 clearly show the importance of active learning, we remained interested in whether the structural information contained in a properly created hypertext organizer could help users learn the structure of a website. In Experiment 2, therefore, two new conditions were created which differ in the amount of structural cues given while still requiring the participant to actively search the website for each answer.

3. Experiment 2

The map from Experiment 1 was replaced in Experiment 2 by a simpler organizer placed in a box on each animal page. In the explicit condition, as shown in Figure 6, the terms visible in the boxes at the bottom of the page identify the location of the current animal page within all three hierarchies of the website. Like the map in Experiment 1, this organizer gives explicit information about the structure of the three hierarchies in the website, and provides an external representation of the user's location in the website. Unlike the map in Experiment 1, this organizer integrated into the animal page, and this organizer is navigable; that is, the terms in the organizer are links to the pages they name. Most importantly, this organizer does not give the location of the next answer.

Figure 6 also shows the implicit version of the animal page, in which the same terms are listed but the terms are now in a single alphabetical list to the right of the animal page. Both versions of the website have the same navigation options, but the explicit version presents those options in a format that gives information about the structure of the website.

The change from a non-navigable map to navigable organizers was made for two reasons. First, this organizer was chosen to be more ecologically valid and less frustrating than the non-navigable map used in Experiment 1. It also allows participants who know the location of the next animal to jump straight to the top of any of the three hierarchies, rather than by moving to the top of the hierarchy one page at a time. Participants performing a more exhaustive search of one hierarchy will likely not take advantage of this ability, since they can start their search at the nearest group page. This change should therefore selectively lower the number of pages a knowledgeable participant searches, helping to distinguish such a participant from others who may not know where they are going.

The same task was performed in Experiment 2 as in Experiment 1, except that the questions were presented on computer rather than in a paper booklet. The same dependent measures were collected in this experiment as well.

In addition, three individual differences measures were given: a short background knowledge test, and two spatial tests, the Paper Folding test and a modified version of Baddeley's Three-Minute Test of Intelligence. Prior research has suggested that both

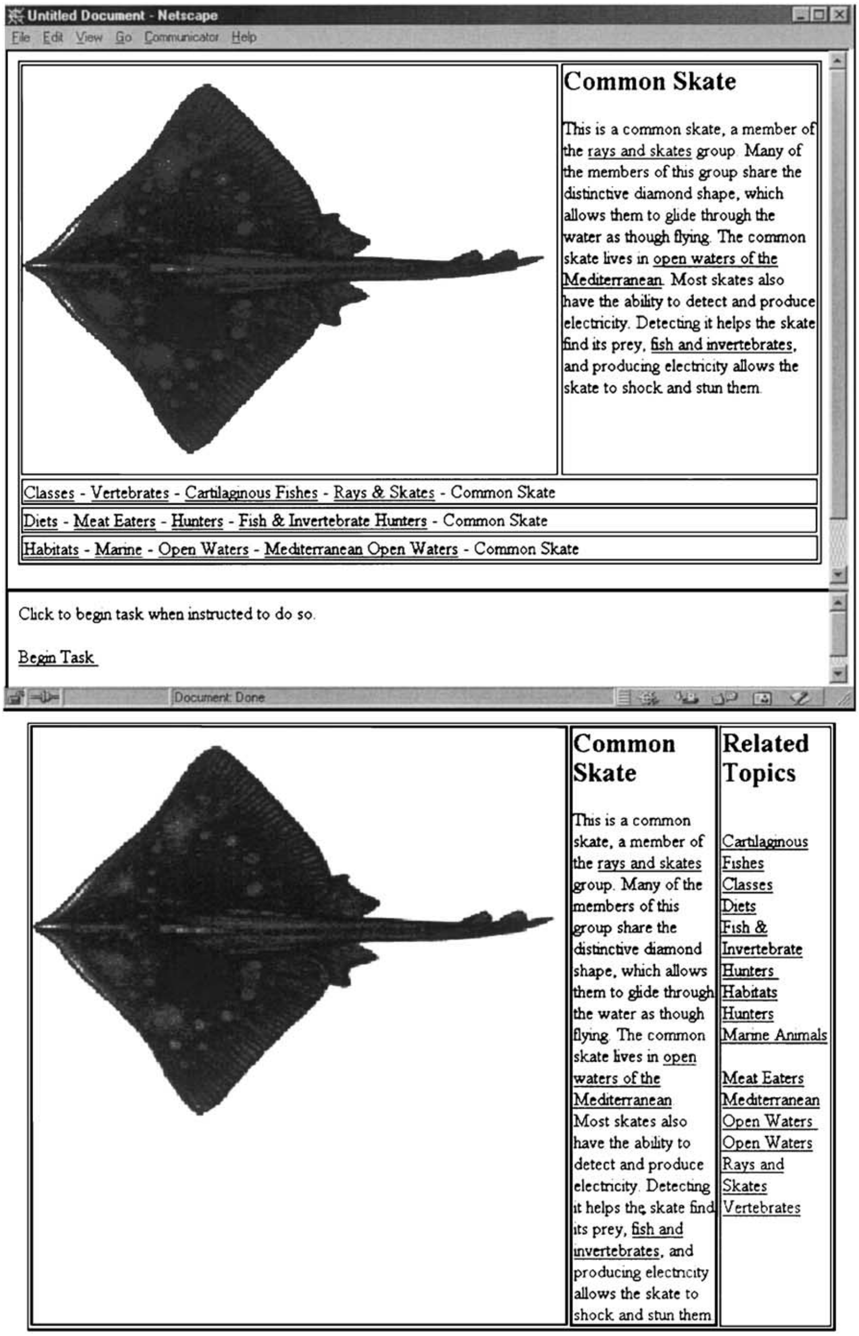

FIGURE 6. Two versions of the common skate page from Experiment 2; top version is from the explicit condition; bottom version is from the implicit condition. The explicit version is shown in the frames as they appeared to participants in Experiment 2.

prior knowledge (McDonald & Stevenson, 1998*a*, *b*) and spatial ability (Vincente & Williges, 1988; Campagnoni & Ehrlich, 1989) predict better performance in hypertext navigation (Chen & Rada, 1996).

Since neither organizer tells the participants exactly what path to follow to get to the next answer, both groups can be expected to be actively involved in the navigation process. This will allow us to focus on whether structural cues presented in a hypertext organizer have benefits for users.

Based on the cognitive load theory, we predicted that the explicit group would be more efficient and accurate across both the learning and test phases, requiring fewer hits and less time across both phases of the experiment. We predicted that the explicit group would perform fewer backups and have lower preference scores than the implicit group. We also predicted that the explicit group would display more structure-related strategies, while the implicit group would display more task-related strategies.

Since neither organizer tells the participant what exact path to follow to the next answer, we predicted that both groups would be cognitively involved in the navigation task. The active learning theory, therefore, does not predict differences between the two groups.

As for the individual difference measures, we predicted that high scores on any of the three measures would lead to more efficient and accurate performance, the use of task-related strategies and lower preference scores. We also predicted that the presence of structural cues in the explicit condition would decrease the differences between the high and low background knowledge participants.

3.1. METHOD

3.1.1. Participants. The participants were undergraduates at the University of California, Santa Barbara. They participated for class credit, and consent was obtained prior to participation. In total, 66 participants completed the experiment, although participants' failure to follow directions and the decision to remove extreme outliers led to the exclusion of 11 of the participants. These participants had not participated in Experiment 1. The mean SAT scores for the explicit group ($M = 1177.54$, s.e. = 26.08) did not differ from that of the implicit group ($M = 1211.82$, s.e. = 26.91), t (44) = -0.92, $p = 0.37$. The mean amount of time spent on the World Wide Web did not differ for the explicit group ($M = 3.65$, s.e. = 0.89) and the implicit group ($M = 5.78$, s.e. = 1.06), t (49) = -1.54, $p = 0.13$. In addition, the ratio of males to females in the explicit group (12/14) was the same as that in the implicit group (8/17), χ^2 (1, $N = 51$) = 1.07, $p = 0.31$. The proportion of participants with web design experience did not differ significantly for the explicit (7/26) and implicit group (6/25), χ^2 (1, $N = 51$) = 0.06, $p = 0.81$. The mean age for the explicit group ($M = 19.08$, s.e. = 0.22) did not differ from that for the implicit group ($M = 18.92$, s.e. = 0.36), t (49) = 0.37, $p = 0.71$. Finally, the mean academic year for the explicit group ($M = 1.54$, s.e. = 0.17) did not differ from that of the implicit group ($M = 1.40$, s.e. = 0.12), t (49) = 0.67, $p = 0.50$.

3.1.2. Materials and apparatus. We created two new versions of the aquatic animals website used in Experiment 1. Figure 6 shows an example page from the explicit and implicit versions of the website. The pictures of the animals and the text, complete with links, were the same as in Experiment 1. In the explicit version of the website, however, the three boxes beneath the picture and paragraph (one box for each hierarchy) contained a list of the groups in that hierarchy that the animal belongs to, in order from most superordinate on down. The listing for each group was also a link to that group's page. In the implicit version of the animal page, one alphabetical list of all the links appeared in a box to the right of the text box. The group pages in Experiment 2 were identical to those in Experiment 1. Thus, the link structure of the two versions was the same.

The same 30 questions from Experiment 1 were used. In the second experiment, however, these questions appeared in a separate frame at the bottom of the screen, with a link to click on to move to the next question. The participants were given an answer sheet with 30 numbered lines on which to write their answers.

The pre-experimental questionnaire was the same as in Experiment 1 except for an additional section consisting of five questions about animals in the website. An example question is "How are leeches used in modern medicine?" These questions were similar to those used in the experimental task. The post-experimental questionnaire was the same as that used in Experiment 1.

In addition, the paper folding test (Ekstrom, French & Harmon, 1976) was administered. This test presents a series of line drawings showing a sheet of paper being folded up, and a hole punched through the paper. The participant then has to choose the correct drawing of the unfolded sheet of paper from five alternatives. We also administered a modified version of Baddeley's three-minute test (Baddeley, 1968). Participants read 32 sentences describing the spatial relationship of two symbols, a star and a plus, compared the sentence to the picture to the right of the sentence, and then indicated whether the sentence is true or false by circling a "T" or "F." A line of the test is

"T F The star is to the right of the plus, (15 ×) ❋ (12 ×) + ."

Both tests were presented in a paper-and-pencil format.

As with Experiment 1, Experiment 2 was run using Netscape Navigator on a Macintosh G3 or PowerMac 7600 computer system, each using a 17 in monitor.

3.1.3. Procedure. Upon entering the room, participants were randomly assigned to either the map or no-map condition and tested in groups of 1–3 per session. They were seated in cubicles to minimize inter-participant interference.

The pre-experimental question was administered as in Experiment 1. The participants were then given the instructions for the experiment, which highlighted the links and the hierarchies. In particular, participants in the explicit group were told to pay attention to the order of the terms in the organizer at the bottom of the screen, and that the order of the terms represented the structure of the links in the website.

The experiment then continued as in Experiment 1, until the participants had finished the usability questionnaire. After this, the paper folding test was administered, with a 3-min time limit. Then the modified Baddleley's three-minute test was administered. This version was only half as long as the original, so participants were given 90 s to finish it. The entire experiment took between 90 and 120 min.

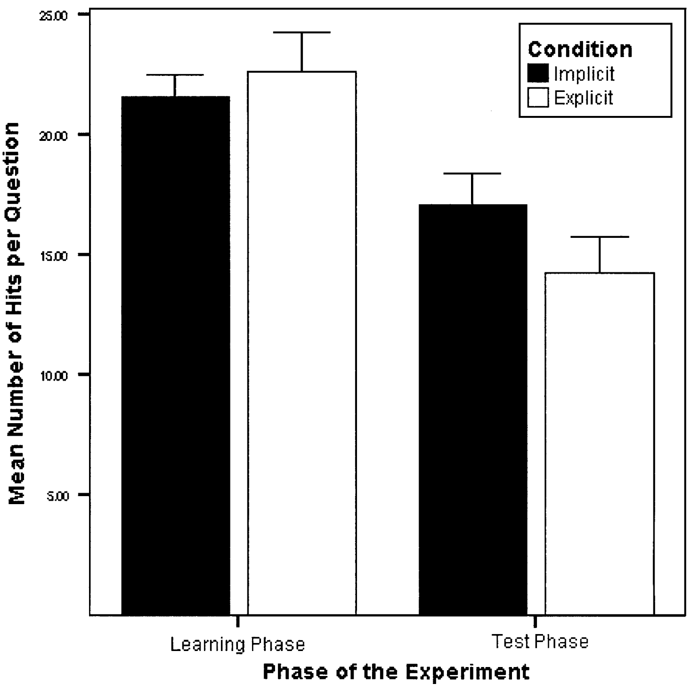

FIGURE 7. Mean number of hits per question (+S.E.) as a function of the phase of the experiment and treatment condition—Experiment 2.

3.2. RESULTS

Once again, the access log generated by the network web server was used to compute the statistics below.

3.2.1. Efficiency of performance. The mean number of hits per question and the mean time per question were computed as in Experiment 1, producing two mean hits scores and two mean times, one for questions 2–20 (learning phase), and one for questions 21–30 (test phase).

Based on the cognitive load hypothesis, we predicted that the explicit group would have a lower mean number of hits in both the learning and test phases. As seen in Figure 7, the explicit group ($M = 18.45$, S.E. = 1.00) did not significantly outperform the implicit group ($M = 19.31$, S.E. = 1.02), $F(1{,}49) = 0.37$, M.S.E. = 18.87, $p = 0.55$, and there was no significant interaction between the treatment condition and the phase of the experiment, $F(1{,}49) = 2.10$, M.S.E. = 95.43, $p = 0.15$. There was a significant decrease in the mean hits per question from the learning phase ($M = 22.07$, S.E. = 0.96) to the test phase ($M = 15.68$, S.E. = 1.00), $F(1{,}49) = 22.89$, M.S.E. = 1042.00, $p < 0.01$. The presence or absence of the explicit structural information had no significant effect on the number of hits required for participants to find the answer to the questions.

The mean time, in seconds, that participants took to complete the questions was measured for each phase. While the explicit group ($M = 93.42$, S.E. = 4.39) was not significantly faster than the implicit group ($M = 99.32$, S.E. = 4.47) , $F(1{,}49) = 0.89$,

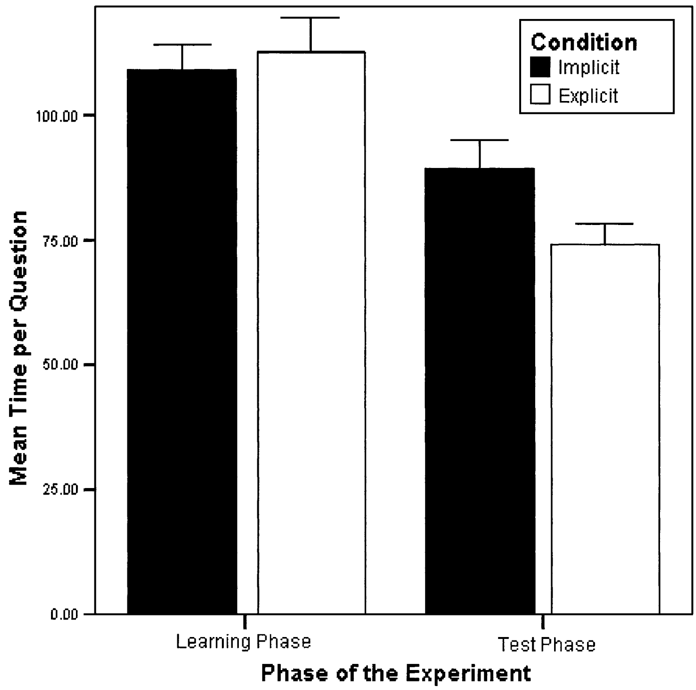

FIGURE 8. Mean time per question (+S.E.), in seconds, as a function of the phase of the experiment and treatment condition—Experiment 2.

M.S.E. = 889.90, $p = 0.35$, an interaction was found between the phase of the experiment and the condition, $F(1,49) = 4.05$, M.S.E. = 2292.84, $p = 0.05$. We used simple main effects testing to determine whether the interaction was that predicted by the cognitive load hypothesis. As can be seen in Figure 8, during the learning phase the explicit group ($M = 112.71$, S.E. = 6.10) was not significantly faster than the implicit group ($M = 109.13$, S.E. = 6.22), $F(1,49) = 0.169$, M.S.E. = 162.95, $p = 0.68$. During the test phase, however, the explicit group ($M = 74.12$, S.E. = 4.80) was faster than the implicit group ($M = 89.51$, S.E. = 4.89), $F(1,49) = 5.04$, M.S.E. = 3019.79, $p<0.05$. This gives some support to the cognitive load hypothesis, since the organizer giving explicit information about the website's structure helped participants become faster.

3.2.2. Accuracy of performance. The number of times the participant backed up during the last 10 questions was also recorded, as in Experiment 1. The cognitive load hypothesis predicted that the explicit group would back up less often than the implicit group. However, the number of backups for the explicit group ($M = 34.77$, S.E. = 4.42) was not significantly different from the number of backups for the implicit group ($M = 40.71$, S.E. = 4.46), $t(48) = -0.945$, $p = 0.35$. This does not support the cognitive load hypothesis.

3.2.3. Quality of models. A preference score was calculated as in Experiment 1. This score represents the degree to which a participant chooses to use one hierarchy more than the rest. A high score indicates a participant unwilling to change hierarchies even when cues would suggest a shorter path in another hierarchy. A low score indicates a participant who uses all three hierarchies approximately equally. A high score is theorized to result from a participant internally reducing the size of the website to make it easier to handle, a sign of an incomplete mental model. The mean preference score for the explicit group ($M = 59$, s.d. = 33) was significantly higher than that for the implicit group ($M = 35$, s.d. = 22), $t(49) = 3.09$, $p < 0.01$. Thus, participants who were given explicit information showed a greater tendency to use the same one or two hierarchies over and over, rather than using all three more equally.

3.2.4. Strategy use. As in Experiment 1, the participants' choices of strategies were recorded. In this experiment, however, the structural strategies were more diverse. Namely, participants who took advantage of the ability to "skip ahead" in their path to the next animal were labeled as using structural strategies. One participant, however, produced results that made it impossible to characterize that individual as using either type of strategy more often, and was removed for this analysis only. The cognitive load hypothesis predicted that the explicit group would be more likely to use such structural strategies, whereas those not given organizers would be more likely to use task strategies. The explicit group was no more likely to use task strategies (12 out of 26) than the implicit group (13 out of 24), χ^2 (1, $N = 50$) = 3.21, $p = 0.57$. Giving participants explicit information about the structure of the website had no effect on their strategy choices. This pattern of results does not support the cognitive load hypothesis.

3.2.5. Self-reports of usability. The cognitive load theory predicted that the explicit group would rate the website as more usable than the implicit group. The response to the statement "I think I know how the information on this website is organized" for the explicit group ($M = 3.69$, s.e. = 0.14) was no different than for the implicit group ($M = 3.84$, s.e. = 0.11), t (49) = −0.807, $p = 0.42$. The mean response to the statement "I could use this website to find out more information about other fish" was also no different for the map group ($M = 4.08$, s.e. = 0.19) than for the no-map group ($M = 4.04$, s.e. = 0.13), t (49) = −0.812, $p = 0.86$.

3.2.6. Individual differences. Three measures relating to individual differences among participants were administered: the paper folding test, a modified version of Baddeley's three-minute intelligence test, and the domain knowledge pretest. For each test, participants were split into low- and high-ability groups based on a median split.

Based on the existing literature, we predicted that the high ability group on the paper folding test would be more efficient in navigating the website. While the high ability group ($M = 17.84$, s.e. = 1.09) did not require significantly fewer hits than the low ability group ($M = 19.68$, s.e. = 0.99), $F(1,47) = 1.55$, m.s.e. = 80.96, $p = 0.22$, a marginally significant interaction was found between phase and paper folding score, $F(1,47) = 4.02$, m.s.e. = 173.25, $p = 0.05$. As can be seen in Figure 9, a simple main effects test found that the high ability group ($M = 22.63$, s.e. = 1.48) did not require significantly fewer hits than the low ability group ($M = 21.78$, s.e. = 1.34) during the learning phase, $F(1,47) = 0.180$, m.s.e. = 8.67, $p = 0.48$. During the test phase,

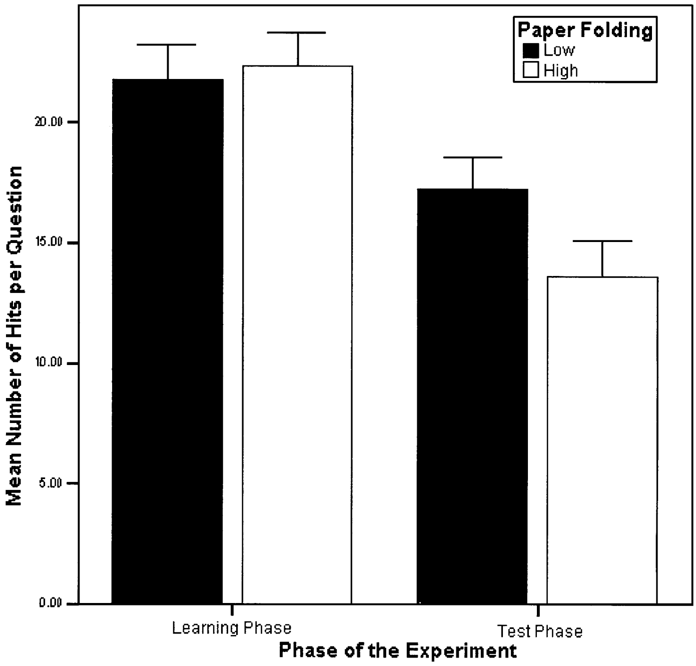

FIGURE 9. Mean number of hits per question (+S.E.) as a function of the phase of the experiment and paper folding score—Experiment 2.

however, the high ability group (M = 13.05, S.E. = 1.47) required fewer hits than the low ability group (M = 17.57, S.E. = 1.33), $F(1,47) = 5.22$, M.S.E. = 245.53, $p < 0.05$. There was no interaction predicted between paper folding ability and organizer condition and none was found, $F(1,47) = 0.038$, M.S.E. = 2.00, $p = 0.92$. Participants who were good at manipulating visual imagery learned to navigate more efficiently.

Based on the same literature we predicted that the high ability group on Baddeley's three-minute test would have a lower mean hits per question than the low ability group. The difference in mean number of hits per question between the high ability group (M = 17.33, S.E. = 0.99) and the low ability group (M = 19.96, S.E. = 1.02) was found to be marginally significant, $F(1,47) = 3.43$, M.S.E. = 165.11, $p = 0.07$. This difference can be seen in Figure 10. There was no interaction of Baddeley's test score and phase of the experiment predicted and none was found. $F(1,47) = 1.96$, M.S.E. = 87.95, $p = 0.17$. Likewise, no interaction between Baddeley's test score and organizer condition was predicted and none was found, $F(1,47) = 2.07$, M.S.E. = 99.54, $p = 0.16$. Participants who were better at transforming verbal descriptions into mental images did better at navigating the website from the very beginning.

As mentioned above, domain knowledge has also been linked to hypertext navigation. We therefore predicted that participants scoring higher on the pretest would have a lower mean number of hits per question. However, high domain knowledge participants (M = 17.94, S.E. = 0.91) did not require significantly fewer hits than low domain knowledge participants, (M = 19.92, S.E. = 1.05), $F(1,47) = 2.06$,

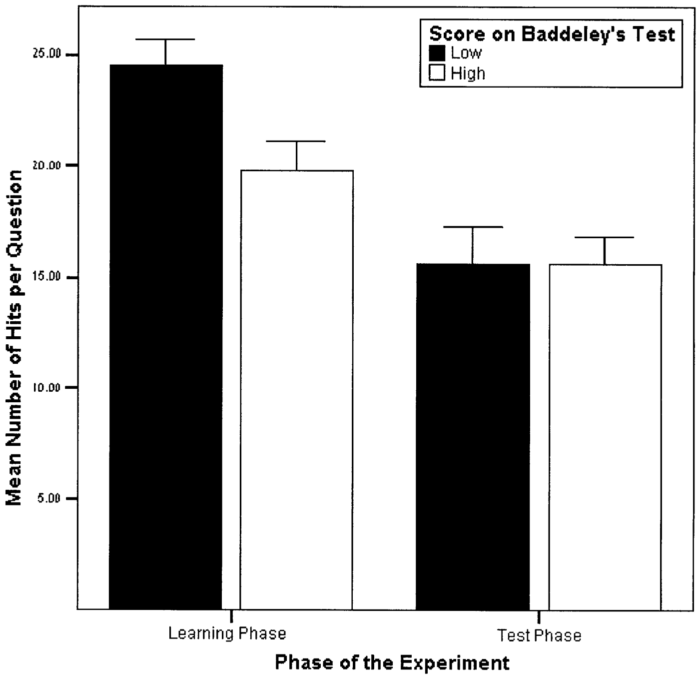

FIGURE 10. Mean number of hits per question (+S.E.) as a function of the phase of the experiment and score on the modified Baddeley's three-minute reasoning test—Experiment 2.

M.S.E. = 98.16, $p = 0.16$. A marginally significant interaction between domain knowledge, treatment condition, and phase was found, $F(1,47) = 3.81$, M.S.E. = 157.53, $p = 0.06$. Using a simple main effects test, we found that when given the explicit organizer, the high knowledge participants ($M = 9.93$, S.E. = 1.7) required fewer hits than the low knowledge participants ($M = 19.36$, S.E. = 1.83) in the test phase only, $F(1,47) = 14.23$, M.S.E. = 574.56, $p < 0.02$. This finding, which can be seen in Figure 11, contrasts not only with our predictions but also with past research showing that organizers decrease the difference between high and low prior knowledge participants. If the information signaling the structure of the website is expected to help low prior knowledge participants, then the explicit organizer should have been more useful than the implicit.

We also tested the effects of these three individual differences measures on the time required to answer each question. We found the same pattern of results as for the hits required data described above.

3.3. DISCUSSION

3.3.1. Summary of results. In Experiment 2, the different organizers did not affect the number of pages participants searched in either phase of the experiment. On the time data, however, a significant interaction was found, with the explicit group outperforming the implicit group during the test phase. The two groups were not different

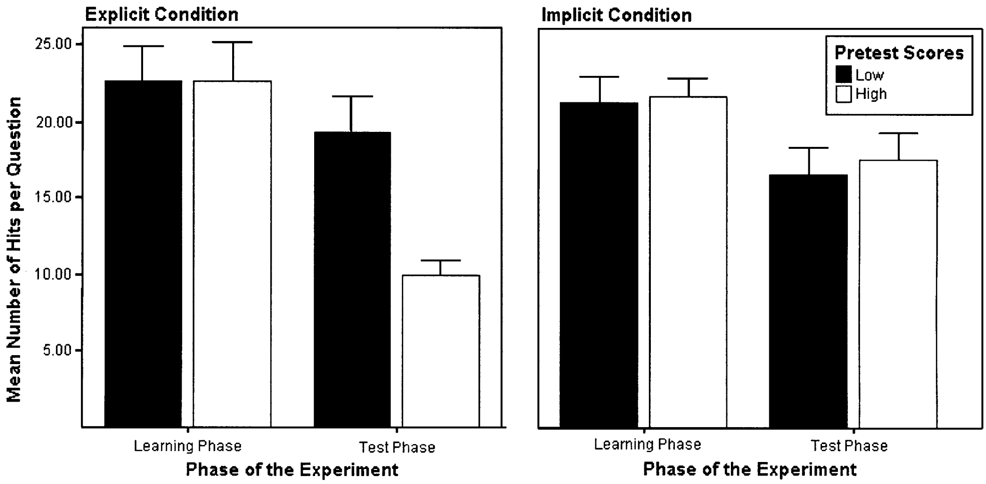

FIGURE 11. Mean number of hits per question (+S.E.) as a function of the treatment condition, phase of the experiment, and domain prior knowledge pretest—Experiment 2.

in their strategy choices, although the explicit group tended to focus on one or two hierarchies only. The results therefore give partial support to the cognitive load theory.

In addition, individual differences were found to affect performance. Participants with high scores on the paper folding test (Ekstrom *et al.*, 1976), while initially showing no advantage in efficiency, outperformed those with lower scores during the test phase. This supports the view that the ability to manipulate visual images is useful for learning to navigate a hypertext document, and suggests that the effects found in past studies may have been due to differences between ability groups towards the end of, rather than throughout, the experiment. Participants with high scores on the modified version of Baddeley's three-minute test were marginally more efficient than those with lower scores. This suggests that the ability to create spatial images from verbal descriptions is useful in all stages of navigating a hypertext document. Finally, participants who scored well on our test of prior knowledge were helped more by the explicit organizer in the test phase only, which is the opposite of what the past literature suggests. This is likely an indication that our prior knowledge measure was flawed.

These results suggest that explicit structural information contained in an organizer does help users become more efficient in their navigation of a hypertext document over time. Since the same information could be gleaned from the implicit organizer with effort, this supports the cognitive load theory. Lowering the effort required to comprehend the structure of the website improved user navigation over time.

4. General discussion

These results are inconsistent with the idea that graphical organizers are, without qualification, helpful for users. The picture that emerges is more complex. While graphical organizers can have benefits for users, these benefits are not without costs.

These results suggest a tradeoff between organizers that are initially useful and those that are useful for less novice users. The map given in Experiment 1 was helpful for participants, but did not lead to learning of the website structure or later increases in efficiency. It seemed to lead to stereotyped searches with little planning. The primary usefulness of such a map seems to be that it tells you how to get to the next answer.

The data also suggest that it is possible to make the task of navigation too easy. The map group in Experiment 1 did not spend time integrating the information they were gaining while moving through the website, because it was unnecessary to do so. Thus, they never really built up a mental representation of the hypertext document, and were unable to improve their efficiency over time.

The success of the no-map group supports the theory that active learning leads to better performance on a navigation task. These participants were able to adapt to a difficult task and to use the sparse information contained in the names of the links to develop clever search strategies. While past studies have shown that forcing hypertext readers to think about the structure of the document leads to better performance on essay tests (Jonassen & Wang, 1993; Shapiro, 1998), it was by no means certain that they would be better at navigation.

In the case of an organizer that gives information about the structure of the document without telling you how to get to the next answer (as in Experiment 2), no initial advantage is seen. Navigation remains a challenging task for the user. Over time, however, this type of organizer leads to improvements in navigation. This supports the theory that lowering the cognitive load of the navigation task can help the user learn the structure of the website.

4.1. APPLICATIONS FOR DESIGN

These results suggest that the desitgn of a hypertext document depends on the intended users and their goals. A hypertext document built for one-time users seeking specific information would do well to include a map like that in Experiment 1. Such an organizer would maximize the efficiency of a single search, but would not benefit future searches. An example of this would be a customer support website, designed to answer specific questions for users. In such a case, the user is not expected to learn the organization of the website, but merely to find one specific piece of information.

If, however, the goal for users is to become very fast at locating information, and an initial learning period is acceptable, then an organizer like that used in the explicit condition in Experiment 2 is better. An example of this case would be an intranet used by telephone customer service representatives. While their initial searches, during a training period, would be slower, their learning of the document's structure would lead to greater efficiency later.

Finally, if the user's ultimate goal is to learn relationships between concepts (i.e. relationships represented by the link structure of the document), then an organizer like the explicit organizer in Experiment 2 seems to facilitate this type of knowledge. While a direct comparison between the map in Experiment 1 and the explicit organizer in Experiment 2 is not possible, a map like the one used in Experiment 1 does not seem to encourage the user to learn the organization of the website. Examples of such documents would include educational hypertext documents, in which students are expected to integrate the different pieces of information into an overall knowledge structure.

4.2. METHODOLOGICAL IMPLICATIONS

This research indicates that the usability of a graphical organizer changes over time, even a time interval shorter than 2 h. Without the distinction between the learning and test phases, these patterns would not have been visible. When studying the usability of a system given to novice users, it is important to look at changes in behavior over even short periods of time. Users develop mental models rather quickly, and this must be addressed in the design of usability studies.

Second, the use of the strategy measure in both experiments suggests that usability measures need not be limited to speed and accuracy. Significant differences in user strategies can be identified, and these can be related both to software design and to more traditional usability measures. Such measures may be particularly useful for evaluating educational software, as they can be sensitive to the differences between strategic activity based on a developing mental model and stereotyped searches. They may also represent a level of task engagement, which has been theorized to be an important aspect of usability (Norman, 1986).

Finally, while the cognitive load of navigation was theorized to be lower with a more informative organizer, the actual level of cognitive load was never actually tested using a dual-task methodology. Without this step, any conclusions about the influence of cognitive load on hypertext navigation remain tentative.

References

BADDELEY, A. D. (1968). A 3-min reasoning test based on grammatical transformation. *Psychonomic Science*, **10,** 341–342.

CAMPAGNONI, F. R. & EHRLICH, K. (1989). Information retrieval using a hypertext-based help system. *ACM Transactions on Information Systems*, **7,** 271–291.

CHEN, C. & RADA, R. (1996). Interacting with hypertext: a meta-analysis of experimental studies. *Human–Computer Interaction,* **11,** 125–156.

CONKLIN, J. (1987) Hypertext: an introduction and survey. *IEEE Computer*, **20,** 105–109.

DEE-LUCAS, D. & LARKIN, J. H. (1995). Learning from electronic texts: effects of interactive overviews for information access. *Cognition and Instruction*, **13,** 431–468.

EKSTROM, R. B., FRENCH, J. W. & HARMON, H. H. (1976). *Kit of Factor-referenced Cognitive Tests*. Princeton, NJ: Educational Testing Service.

HAMMOND, N. & ALLINSON, L. (1989). Extending hypertext for learning: an investigation of access and guidance tools. In A. SUTCLIFFE & L. MACAULAY, Eds. *People and Computers*, Vol. V, pp. 293–304. Cambridge, UK: Cambridge University Press.

JONASSEN, D. & WANG, S. (1993). Acquiring structural knowledge from semantically structured hypertext. *Journal of Computer-Based Instruction*, **20,** 1–8.

MANNES, S. M. & KINTSCH, W. (1987). Knowledge organization and text organization. *Cognition and Instruction*, **4,** 91–115.

MCDONALD, S. & STEVENSON, R. J. (1998*a*). Effects of text structure and prior knowledge of the learner on navigation in hypertext. *Human Factors*, **40,** 18–27.

MCDONALD, S. & STEVENSON, R. J. (1998*b*). Navigation in hyperspace: an evaluation of the effects of navigational tools and subject matter expertise on browsing and information retrieval in hypertext. *Interacting with Computers*, **10,** 129–142.

MONK, A. F., WALSH, P. & DIX, A. J. (1988). A comparison of hypertext, scrolling and folding as mechanisms for program browsing. In D. M. JONES & R. WINDER, Eds. *People and Computers,* Vol. IV, pp. 421–435. Cambridge, UK: Cambridge University Press.

NICKERSON, R. S. & LANDAUER, T. K. (1997). Human–computer interaction: background and issues. In M. HELANDER & T. K. LANDAUER, Eds. *Handbook of Human–Computer Interaction,* (2nd edn), pp. 3–32. Amsterdam: Elsevier Science B.V.

NORMAN, D. A. (1986). Cognitive engineering. In D. A. NORMAN & S. W. DRAPER, Eds. *User Centered Design: New Perspectives on Human–Computer Interaction*, pp. 31–62. HILLSDALE, NJ: Lawrence Erlbaum Associates, Inc.

SHAPIRO, A. (1998). Promoting active learning: the role of system structure in learning from hypertext. *Human–Computer Interaction*, **13,** 1–35.

SIMPSON A. & MCKNIGHT, C. (1990). Navigation in hypertext: structural cues and mental maps. In R. MCALEESE & C. GREEN, Eds. *Hypertext: State of the Art*, pp. 73–83. Oxford, UK: Intellect.

SWELLER, J. (1988). Cognitive load during problem solving: effects on learning. *Cognitive Science*, **12,** 257–285.

TRIPP, S. & ROBY, W. (1990). Orientation and disorientation in a hypertext lexicon. *Journal of Computer-Based Instruction*, **17,** 120–124.

VINCENTE, K. J. & WILLEGES, R. C. (1988). Accommodating individual differences in searching a hierarchical file system. *International Journal of Man–Machine Studies*, **29,** 647–668.

VORA, P. & HELANDER, M. G. (1997). Hypertext and its implications for the internet. In M. HELANDER & T. K. LANDAUER, Eds. *Handbook of Human–Computer Interaction,* (2nd edn.), pp. 877–914. Amsterdam: Elsevier Science B.V.

Paper accepted by Associate Editor, C. McKnight