Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Int. J. Human-Computer Studies 58 (2003) 125–149
International Journal of Human-Computer Studies
www.elsevier.com/locate/ijhcs

Balancing search and retrieval in hypertext: context-specific trade-offs in navigational tool use

Stephania Padovani, Mark Lansdale*

Department of Human Sciences, Loughborough University, Loughborough, Leicestershire, LE11 3TU, UK

Received 1 October 2002; accepted 10 October 2002
Paper accepted for publication by Editor, B.R. Gaines

Abstract

We compare user navigation performance using two hypertext information sites of identical node structure but embedded in different metaphors. The first is based upon the layout of a house and is consistent with Euclidean space. The second represents social links between people for which a spatial metaphor is not apparent. Search for targets within the structures, and the speed of their subsequent retrieval on a second search, is compared in a $2 \times 4 \times 2$ factorial design manipulating: metaphor (spatial or non-spatial); navigation tools (participants have both a site map and bookmark tool, one of these, or no tools at all) and the time pressure under which navigation is carried out (paced or unpaced). A strong main effect is found in which the spatial metaphor produces higher performance under all conditions. Similarly, time pressure has the general effect of trading-off a faster initial search with less efficient retrieval later. However, navigation tool use is highly context dependent and sometimes counter-intuitive: certain conditions show poorer performance with two navigation aids than one. We argue that navigation tools are mediating structures for activities, such as bookmarking and learning the structure of the site, which represent cognitive investment for future retrieval. In this view, user performance is optimized by the balance of two potentially antagonistic conditions. First, the usability of tools and metaphor must free cognitive resources for planning; but also, the difficulty of the task and the need for planning must remain visible to the user. The implications for design are discussed.

Keywords: Navigation; Spatial metaphor; Memorability; www

*Corresponding author. Tel.: +44-1509-263171; fax: +44-1509-610724.
E-mail addresses: S_Padovani2@yahoo.co.uk (S. Padovani), M.W.Lansdale@lboro.ac.uk (M. Lansdale).

1071-5819/03/$ - see front matter © 2002 Published by Elsevier Science Ltd.
PII: S1071-5819(02)00128-3

1. Introduction

It is generally accepted that navigation through complex information spaces such as the World Wide Web (WWW) is problematic. This is seen both in terms of the efficiency of usage, as measured by wasted time and effort and in failure to access required information (McDonald and Stevenson, 1998); and in terms of user dissatisfaction following from this inefficiency and a general sense of disorientation (Nielsen, 1995). This is hardly surprising. The WWW is a very large hypertext structure that is heterogeneous in content and presentation style, and navigation through it is a complex activity in many senses: (1) the *purposes* for navigation may vary, from differing degrees of browsing through to direct search for specified targets (McAleese, 1989); (2) users' search strategies differ according to task and context (Darken and Sibert, 1996); and (3) the navigating tools they use can also vary widely (Wright and Lickorish, 1990). Further, it is clear that these factors are themselves strongly interdependent (Marchionini and Shneiderman, 1988). From a psychological perspective, there are also a range of issues concerning the cognitive skills underlying this behaviour (Paolucci, 1998; Chou and Lin, 1998) and the role of individual differences (Chen and Rada, 1996). The motivation for the volume and diversity of research in this area is clear.

In this paper we wish to add to this cocktail by widening the issue of what is meant by navigation. Consistent with a spatial metaphor, there is a tendency to see the tools of navigation as means to an end — a way of getting from sites A to B, and the studies of these tools are often in terms of the efficiency with which these tools enable that (Chou and Lin, 1998; Dias and Sousa, 1997; Gomes, 1994). This approach could be said to be concerned with the mechanics of navigation, and concerns itself with what navigation tools can do and how usable they are.

However, navigation is more than that. It is also the mediating process through which users learn, explicitly or implicitly, the structure of the information space. Just as using the underground or walking around gives rise to a different experience and learning about a city, so we expect that different ways of navigating information spaces will influence what we learn about them. In an information space such as the WWW, this structure has two functions. First, it is a context in which subsequent browsing or search will take place. The more you know, the more efficient subsequent search will be and specific, known, targets can be found quicker. Second, one might argue that the structure itself is information in its own right about the nature of what is stored in that information space. This type of meta-knowledge can influence less specific information gathering strategies, such as deciding where it would be interesting to browse, or indeed what it would be interesting and possible to browse about. We argue that the mechanics of navigation, however complex that may be, also has implications for what is learned about the structure and content of the information space. The purpose of this paper is to explore this proposition and to consider the implications of it for design.

In the experiments reported below, we asked participants to search one of two complex experimental WWW sites to find a set of specified target items. The interest here was in the nature and efficiency of this search in relation to a number of manipulated variables, including navigation tools available and the time pressure upon the users. We then asked them: (1) to relocate the specified targets in a second search; and (2) to draw a sketchmap laying out the pages in the structure and the links between them. These measures were designed to indicate the efficiency of the second search and the extent of recall for structure.

The research literature on human cognition suggests a number of issues of interest in this approach.

(1) *Transfer-appropriate processing* (*TAP*): A central tenet of the TAP view (Morris et al., 1977) is that what we recall of information is determined by the cognitive processes at initial exposure and their relationship to the processes mediating subsequent attempts at recall. This idea places the importance upon tool use as mediating the nature of activity (rather than the intention to learn) as a significant factor in learning (see also Hyde and Jenkins, 1973). In these experiments, we study the use of two navigation tools. The first, a site map, emphasizes the structural nature of the site as a whole, and its use can be expected to promote recall of structure because it is the structure of the site that the user is manipulating. The second tool, bookmarking, does not. Whilst it is more efficient in returning to a given previously marked page, it does so directly and without reference to structure. When it comes to recalling structure, this tool can therefore be expected to produce poorer recall. In this study, we are interested in the magnitude of this effect and its impact upon users tool use and search strategies.

(2) *Effort after meaning, planning and time pressure*: It is generally true that the greater engagement in the processing of material, the stronger the learning. For example, Baddeley (1963) found that the harder the solved anagram, the easier it was to remember. If we are designing web sites to promote learning of structure, this leads to the counterintuitive hypothesis that it might pay to make navigation harder. One element in the allocation of effort during learning and performance lies in the notion of planning. It is well known that users are poor judges of when to commit immediate resources to planning activities which might actually save time later, particularly if the immediate payoff is seen as disadvantageous (e.g. O'Hara and Payne, 1998). This is also seen in information management, where the ease of retrieval is largely dictated by the effort put into initial storage (e.g. see Malone, 1983; Lansdale and Edmonds, 1992). In this study, adding conditions in which participants are paced to go quickly can be expected to work against the benefits of effort and planning in the initial stages of the task; enabling some empirical assessment of the impact of this upon performance. The relative ease of use of the navigation tools and the cognitive load placed by different task metaphors are also relevant in this respect.

(3) *Context, concreteness and cognitive overload*: A huge literature in human learning testifies to the benefits of concreteness and a familiar context in learning. There are a number of factors in this. First, a known external structure may offer a range of opportunities for external cues which allow the user to monitor where they are in a memorised information structure and to organize their knowledge appropriately (Bower et al., 1969). Second, familiar data structures allow users to 'chunk' (Miller, 1956; Chase and Ericsson, 1981) information in such a way as to

retain more whilst reducing concurrent cognitive load (an issue held to be highly significant to navigation — e.g. see Conklin (1987) or Kim and Hirtle (1995)). Third specificity of context is held to be significant in the recall of information (e.g. Godden and Baddeley, 1975).

(4) *Spatiality*: A final issue here is the issue of spatiality. "Navigation" as a term has strong spatial connotations, and one might expect difficulties of navigation to follow from the tendency of hypertext structures to violate the rules of Euclidean space. Others point to the importance of spatial metaphors in design (Hammond and Allinson, 1987; Kim and Hirtle, 1995; Smilowitz, 1996; Kim, 1999). This touches upon the question of whether visuo-spatial processing is somehow privileged in human cognition, and whether the use of spatial representation can exploit specific cognitive facilities to enhance user performance. This is a highly contentious theoretical issue in cognitive science, and not one to be resolved in an applied context such as this. However, it is reasonable to investigate, as a purely empirical matter, the impact of spatial context on user performance.

In these experiments, we manipulate these dimensions of performance by building two experimental WWW sites of identical nodal and link structure; one of which is a concrete, highly familiar, spatial context and the other not. The first represents a virtual house whose three-dimensional structure is preserved consistently. Each page is a room, and links are only possible to adjoining rooms or spaces. The user task in this case was to search for a number of common items; namely flowers, and then return to them later for "virtual watering". The second site represents a social grouping in which pages represent individual people and links represent acquaintance. The user task in this case is to act as a detective following a crime in which individuals may know significant facts or know others who might. Hence, navigation is moving from person to person and targets consist of useful facts known by some individuals.

To summarise, the above theoretical perspectives lead us to consider how incidental recall of the structure of a WWW site following an information search can be manipulated by three factors: the navigation tools available (singly or in combination); whether the user is under time pressure to complete an original search; and the information context around which the site is structured. In the 16 experimental conditions below, we manipulate these factors orthogonally in tasks where users are asked to search a web site for a specified number of targets and then to relocate them in a second search. We then attempt to understand their impact upon the nature of the primary search. Following this, we consider the relationship between those different patterns of search, the efficiency of secondary search and users' consequent knowledge of the web site structure.

2. Experimental method

This research comprised 16 experimental conditions manipulating four variables in a between-subjects design ($2 \times 4 \times 2$) the metaphor used to build the Web site, the type of navigational aid available and time pressure.

2.1. Between-subjects factor 1: metaphor used to build the web site

The Web site used in the experiments contained 29 interconnected information nodes and six additional screens with the instructions to be followed during the experimental task. This structure was used to build two different versions of the Web site, each of which are based on a different metaphor.

On the first version, we used a spatial metaphor in which each hypertext node was a room or an external area in a house and the links corresponded to the doors or passages between rooms. Each hypertext node contained a brief description of the room, information about its location in the house and links to other rooms or external areas of the house.

The second version of the hypertext utilized exactly the same structure as the first but was based on a non-spatial metaphor. In this case, the hypertext nodes were students' Web pages and the links corresponded to their interpersonal relationships. Each hypertext node contained information about the student's status (undergraduate/postgraduate) in the same way that the house Web site provided information about the location of the rooms (downstairs/upstairs). The descriptions of the rooms were substituted by some information about the student's place of birth, personality and main interests. Each student was linked to their friends and neighbours just as each room was linked to the adjacent rooms.

This study focuses upon variations in the representation of site *structure* (i.e. inter-node relationships). To minimize the confounding effects of intra-node representations, the pages within both sites were purely textual without any graphic differentiation among its screens (except for the site maps provided in some experimental conditions).

2.2. Between-subjects factor 2: navigational aid available

Two different types of navigational aids were selected for the experiments: a bookmark tool and a global site map. These navigational aids have completely distinct characteristics in terms of how they function, how they can be accessed and the level of processing involved in their utilisation.

In this case, the *bookmark tool* consisted of a mechanism of storing Web pages provided by the Internet browser itself (in Netscape 3.01 for Macintosh). It worked by selecting the command "add bookmark" (on the "bookmarks" menu) when the user visited an important page and wanted to mark it. The user returned to any bookmarked page by selecting its title from the resulting bookmark list and did not need to leave the Web page being visited in order to read the bookmark list or access a page from it. The user also has the possibility of renaming the pages, modifying the order in which they appear in the list or grouping pages into customized categories. These functions were available in a separate window called bookmark organizer.

The *global site map* is a graphic representation of the whole Web site structure, providing an overview of the information space. It is a tool provided by the Web site itself, rather than by the Internet browser. In the Web sites used in this study, the site map appeared on a separate screen and could be accessed by clicking on a textual link placed at the bottom of every information node. Because of the size of the site map, it required some scrolling to visualize it all. This type of navigational aid allowed the user to go straight to any information node by clicking on its name as a textual link on the map. Thus, it can be regarded as a graphic list of short cuts. Unlike the bookmark tool, the site map provided did not permit any customization. The names and the quantity of rooms appearing on the map were permanent. The only changes occurring during a user's navigation were the colour of the links that changed as a means of signalling already visited nodes; automatically indicating which areas of the site had been visited and which not.

Two different versions of a global site map were produced in accordance with the two metaphors used to build the Web site. In the spatial metaphor, the structure of the Web site was depicted using a two-dimensional house plan. The site map available for the the non-spatial metaphor consisted of an abstract network representation of the Web site in which links represented friendships. The position of the links on the site map screen was approximately the same for both maps.

2.3. Between-subjects factor 3: time pressure

In all the experimental conditions, half of the participants were put under paced time pressure. These participants were instructed to finish the task as fast as they could trying not to spend more than 6 s per screen (apart from the instructions screens where they could take as long as they needed to understand the task they had to perform). Participants in the non-paced conditions were told that they had no time limit to complete the task and that, as well as completing the task, they should try to learn how to move around the Web site.

2.4. Experimental task

During the experiment, participants were required to perform a specific task comprising two different phases; the first of which involved search and the second retrieval. Although the majority of the experimental works found in the pertinent literature limit themselves to searching tasks, we opted for the inclusion of these two types of tasks in order to try to simulate a situation as close as possible to real tasks performed by Internet users. This option is supported by previous empirical research, which has shown that users actually revisit a considerable number of Web pages (e.g. Tauscher and Greenberg (1997) found a recurrence rate of 58%).

The first part of the experimental task (search) involved finding five targets (specific pieces of information contained in five different nodes) and then accessing a specific information node whose title was given to the participants in the instructions. The second part of the task (retrieval) required the participants to return to their starting point or visit a pre-defined node whose title was also given and then alternate the retrieval of the targets found on the first part with visits to those two pre-defined information nodes. The last step consisted of going back to that particular information node where the participants started the second part of the experimental task.

The instructions varied according to the metaphor used in the Web site. Hence, although the participants had to find and retrieve exactly the same hypertext nodes, the "story" behind the experimental task was completely different for each Web site version.

Participants who performed the task based on the spatial metaphor were asked to water five plants inside the house, whereas those whose task was based upon the non-spatial metaphor were required to contact five witnesses for interrogation. The participants received these instructions in the following ways

- Spatial-metaphor Web site: "According to Mrs Robinson's list, your task consists of: finding the five flowers around the house; getting the watering can in the gardening shed; filling the watering can/watering the flowers; returning the watering can to the gardening shed."
- Non-spatial-metaphor Web site: "According to Inspector Smith, your task consists of: finding the five witnesses among the Geography students; registering their names at the Police Station; arranging appointments with the students representatives; contacting the witnesses for the interrogation; taking the witnesses' statements to Inspector Smith at the Police Station."

2.5. Participants

Two hundred and thirty students from Loughborough University served as participants. Ten participants took part in each spatial-metaphor-non-navigational-aid condition and 15 in all the other conditions. Each participant was tested individually and took part only once in only one experimental condition.

2.6. Procedure

The experimental sessions began with the participant receiving generic verbal instructions on what the experiment would involve (navigate through a Web site in order to complete a specific task), the navigation mechanisms available to navigate the Web site (textual links, back and forward buttons and navigational aids when they were available) and the time pressure factor. As far as the navigational aids are concerned, we took some time to show the participant how to access each tool and how it functioned. No previous instructions about the content, structure of the Web site or the kind of task to be performed were given to the participants. All the information about the task they had to do was presented in the instruction screens.

At the beginning of the experimental task, the Web site provided a main instruction screen with all the steps involved on the task. As the participant progressed through the Web site and completed each step, there were additional instruction screens explaining in more detail the next steps of the task.

The experimental sessions took place in a quiet laboratory. The machine and Internet browser used to run the experiments were a Macintosh Performa 475 and Netscape 3.01, respectively. The Web sites were run locally so as to prevent unexpected variations in the speed of access from one experimental session to another. The participants' navigation was automatically video-recorded during the experimental sessions.

Following the experimental task, participants were required to sketch (from memory) a site map for the Web site they had just navigated in as much detail as they could remember. The participants were not told in advance that they would have to sketch a site map after the completion of the experimental task because our interest was to verify how much they would learn about the Web site structure as a result of performing the experimental task. After sketching the site map, each participant completed a questionnaire mainly about memory strategies, navigation strategies and the difficulty to draw the site map.

2.7. Data collection, tabulation and dependent measures

Three forms of data were collected from the experimental sessions: a recording of the users' navigation performance; site map sketches; and finally, completed questionnaires measuring their subjective judgements upon a range of aspects of the task. For the most part, these qualitative data are not analysed here, other than illustrative examples to supplement analysis of the quantitative data.

Each participant's navigation in both phases of the experiment was tabulated as a sequence of visited screens together with: the time spent viewing each screen; the mechanism used to access each screen; and the role of the screen within the experimental task. These navigation tables were used to generate the dependent measures needed to quantitatively analyse the participants' navigation. These included time measures, navigation pattern measures, navigation efficiency measures, utilization of supplementary navigation tools and task accuracy measures. The definitions of these measures are given in Appendix A.

To quantify recall of the Web site structure in terms of recall of the information nodes names and correct placement on the sketch, total sketch accuracy was calculated by the formula:

$$Sk(tot) = \frac{(\text{number of screens correctly recalled} + \text{number of screens correctly placed})}{2(\text{number of screens visited by the user})}$$

This being a device which represents both the recall of individual pages and memory for the relationship between them.

The questionnaires utilized in this research used between 11 and 15 multiple-choice questions and one open-ended question, depending upon the particular experimental condition. It aimed to gather supplementary information about the users' strategies and difficulties that could not be visualized through the experiment and the site map sketches. The first part of the questionnaire focused on the user's selection criteria and memory strategies during the realization of the task. The second part concentrated upon the "feeling of being lost". The third part was about the utilization of the navigational aid and in which aspects such tool actually helped (e.g. sense of orientation, task accuracy, speed of task completion, comprehension of the

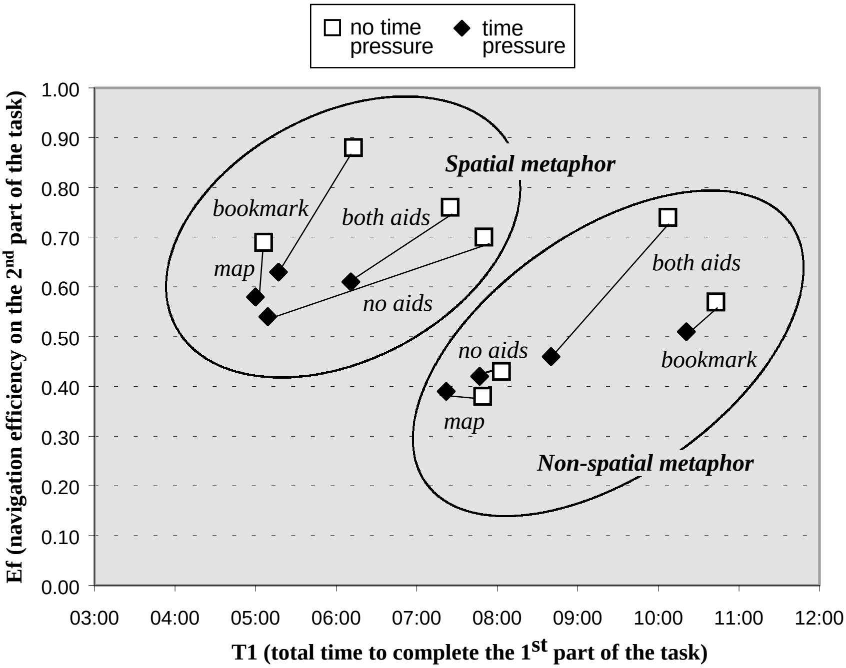

Fig. 1. The relationship between total time to complete the first (search) phase of the task and the efficiency of search in the second (retrieval) stage. Conditions that are common with respect to tool availability and site metaphor and differ only with respect to the presence or absence of paced time pressure are connected for illustrative clarity.

Web site structure). Finally, the users were asked to rate the difficulty faced to sketch the site map and the accuracy of the sketch produced.

3. Results

In the interests of clarity, we present only those analyses which best illustrate the overall pattern of observed results.[1] As an advanced organizer to this, Fig. 1 illustrates the relationship between two variables for each experimental condition: the duration of initial search (T_1) and the efficiency (*Ef*) of the secondary search, which is a composite measure of the number of steps taken to complete the retrieval task and the number of targets successfully recovered (see Appendix A). These capture many of the essential outcomes of this study. The plot is additionally annotated by (1) linking datapoints of equivalent conditions varying only in whether

[1] With 16 experimental conditions and 39 measures each, a comprehensive presentation of all possible data is neither possible nor desirable. Readers are referred to Padovani (2001) for a more comprehensive treatment.

Table 1
General performance measures in the first (search) phase of the task for spatial and non-spatial metaphors.

Measure	Definition	Mean observed values		ANOVA result (all 1df)
		Spatial site	Non-spatial site	
*T*1 (s)	total time to complete phase	361	531	197.78**
*St*1	number of screens visited	50.8	58.1	17.22**
*Nr*1	navigation redundancy	1.72	1.97	14.62**
*Bk*1	proportion of screens accessed by "back" button	0.18	0.26	24.50**
*Tg*1	proportion of targets found	0.96	0.92	9.04**

**$p<0.01$.

users were paced or not during initial search; and (2) encircling datapoints emanating from the conditions run in the spatial context and those in the non-spatial site. Both of these annotations are straightforward in this case because the main effects of time pressure and metaphor are very clear. For clarity, we present the analysis of performance in relation to these two factors first before considering the more complex interactions between the use of navigation tools and these factors.

3.1. The effect of spatial and non-spatial task context

Table 1 summarizes the main effects of context upon general performance measures in the first stage of the task. All are highly significant and represent a considerable advantage to users of a familiar spatial context. Thus they are generally faster to complete the search; access fewer screens; access more targets; make fewer redundant steps; and use the "back" button on a lower proportion of trials.[2] They are also more likely in the spatial context to report carrying out a selective, rather than exhaustive, search of the site [$2I$ (1df, $N=218$) = 110.339 ($p<0.005$)]; and more likely to report having used explicit strategies to keep abreast of the number of targets found and those still outstanding [$2I$ (3df, $N=230$) = 17.373 ($p<0.005$)].

The pattern in the second stage of the task is equally clear (see Table 2). In the spatial context, users are quicker to complete the search; use fewer steps to do so; are less likely to visit unnecessary screens on the way; access more targets; and are generally more efficient in search. They are also more likely to report having used explicit cognitive strategies to manage the process of keeping track of how many targets have been recovered. Finally, the test of memory for the structure of the two sites reveals a very large advantage for users recalling the spatial context [*Sk*(*tot*) for

[2] We report throughout simple contrasts. It would have been technically feasible to carry out a comprehensive ANOVA of all possible main effects and interactions simultaneously. We doubt, however, that the interpretation of such an analysis would have aided the clarity of presentation in what is essentially an investigative study.

the spatial metaphor = 0.54, for the non-spatial metaphor = 0.18. $F(1,214) = 468.237$ ($p < 0.01$)].

To summarize, it is very clear that the familiar spatial context confers very large advantages over a non-spatial equivalent structure. In performance terms, users are considerably more efficient in both search and retrieval. Further, as indicated both by the likelihood of using selective (as opposed to exhaustive) search strategies, and strategies to enable them to keep account of how many and which targets have been found (in the first stage) and retrieved (in the second), it is reasonable to assume that the easier spatial task context allows the users to devote cognitive capacity to the strategic planning of the task. None of these conclusions are unexpected to cognitive psychologists, although the magnitude of some of these effects is larger here than is often observed.

3.2. The effects of time pressure upon performance

Comparison of the eight pairs of conditions in Fig. 1, contrasting users under time pressure and those not, suggests that time pressure always produces faster initial searches and generally less efficient subsequent search. For the first phase of the task, Table 3 illustrates that time of initial search is significantly faster in time-pressured users. However, comparisons of the number of steps taken; the redundancy of search; the use of "back" buttons and the numbers of targets found all produce insignificant changes. This does not mean that the effects of instructions upon search is just to make users go faster. Specific comparisons in the use of navigating tools indicates, for example, that users under time pressure bookmarked fewer targets [an average of 3.1 as opposed to 4.6 in the non-paced conditions $F(1,112) = 15.583$ ($p < 0.01$)] and viewed more screens before using the map [13.4 screens vs. 9.3 $F(1, 112) = 8.062$ ($p < 0.01$)] when available.

The effects of that trade-off and increased speed of processing are seen clearly in the retrieval stage (see Table 4). Here, despite instructions to be as quick as possible, total search times are increased; as are the number of steps taken and the number of unnecessary screens visited. Accordingly, the efficiency of search is reduced. Further,

Table 2
General performance measures in the second (retrieval) phase of the task for spatial and non-spatial metaphors

Measure	Definition	Mean observed values		ANOVA result (all 1df)
		Spatial site	Non-spatial site	
*T*1 (s)	total time to complete phase	183	270	65.36**
*St*2	number of screens visited	36.2	42.5	10.92**
Ef	navigation efficiency	0.67	0.45	46.53**
USr	unnecessary screens revisited	0.47	0.64	5.17*
*Tg*1	proportion of targets reaccessed	0.94	0.77	80.24**

*$p < 0.05$, ** $p < 0.01$

Table 3
General performance measures in the first search phase of the task for paced and not-paced performance

Measure	Definition	Mean observed values		ANOVA result (all 1df)
		Not paced	Paced	
*T*1 (s)	total time to complete search phase	474	418	21.75**
*St*1	number of screens visited	55.0	53.9	0.4
*Nr*1	navigation redundancy	1.82	1.86	0.49
*Bk*1	proportion of screens accessed by "back" button	0.21	0.23	1.14
*Tg*1	proportion of targets found	0.95	0.93	1.65

**$p<0.01$.

Table 4
General performance measures in the second (retrieval) phase of the task for paced and not-paced performance

Measure	Definition	Mean observed values		ANOVA result
		Not paced	Paced	(all 1df)
*T*2 (s)	total time to complete retrieval phase	211	242	8.39**
*St*2	number of screens visited	37.1	41.6	5.40*
Ef	navigation efficiency	0.62	0.50	18.41**
Usr	unnecessary screens revisited	0.43	0.67	11.00**
*Tg*1	proportion of targets reaccessed	0.89	0.82	14.23**
Sk (*tot*)	accuracy of final sketch	0.38	0.34	6.01*

*$p<0.05$, **$p<0.01$.

the number of targets retrieved in the time-pressured conditions is fewer and the users' memory for the two sites is impaired.

These results are easily interpreted in terms of two processes. First, the effects of time pressure on the users is to reduce the likelihood of their employing navigating aids. In the case of bookmarks, this has the obvious outcome in retrieval that a higher proportion of targets have to be re-found in the retrieval phase of the task by search rather than by direct access. A similar argument applies to the use of site map, where, particularly in the spatial metaphor condition, users regularly reported remembering target pages in terms of their location on the map, making the map equivalent to bookmarks insofar as the users can go directly to targets instead of searching for them. Effort invested in these activities — effectively planning for the retrieval stage — is being sacrificed for extra speed.

Apart from influencing the use of navigation aids (i.e. the users' search strategies), it appears that pacing the users under time pressure had a small but significant effect upon what they learn of the structure of the site. In an extended search of complex sites, such as here, such learning might be expected both as a conscious strategy on the part of the user to facilitate the search, and also as a by-product of that search process because the activity of thinking about where to go next is precisely the kind of processing which generates recall (see Lane and Robertson (1979) for a similar argument in remembering the layout of a chessboard). The reduction of recall scores in the time-pressured condition suggests that users in this condition spend less mental effort thinking about the structural arrangements of the pages.

3.3. Issues in the use of navigation aids

Whereas the effects of time pressure and task context are large and mostly transparent, performance with the use of different navigational aids is complex for two different reasons. First, we have to consider the effect of these aids on both searching and retrieval performance, and then on the interaction of one upon the other. Second, the observed patterns appear to be context-specific. For example, looking at the non-spatial context data illustrated in Fig. 1, two aids produce more efficient retrieval than one; as Hammond and Allinson (1989) would predict. However, when we look at the spatial metaphor conditions, this is no longer the case, with use of the bookmark alone producing more efficient retrieval than when both aids are available. We need to explain why the pattern observed in the one context is reversed in the other.

The first thing to note is that these navigation aids are, in themselves, effective. Considering the site map first, the availability of this tool allows the user to jump around the site rather than traversing page by page; hence reducing the number of pages accessed. Thus, for example, there is a significant reduction in the number of pages searched in the first stage of the task when the map is present compared to the unaided condition [an average of 47.2 with the map present and 65.1 when not $F(1, 108) = 35.576$ ($p < 0.01$)]. Equally, the bookmark tool is highly effective at reducing the number of pages visited in the second stage of the task, since each target bookmarked requires only one selection. All bookmarking conditions show a significant advantage in this respect over the no-aids conditions [an average of 57.6 pages visited with no-aids compared to 28.8 with the bookmark available $F(1, 108) = 92.596$ ($p < 0.01$)].

Table 5
A comparison of basic performance measures with and without the site map

Site metaphor	Pacing condition	Differences from no-aids condition				Absolute measures	
		$T1$ (s)	$St1$	$Nr1$	M_Pr1	Ef	$Tg2$
Spatial	Unpaced	164*	32.2*	1.23*	0.39	0.69	1.00
	Paced	9	12.1*	0.75*	0.28	0.58	0.88
Non-spatial	Unpaced	14	19.0*	0.86*	0.21	0.38	0.76
	Paced	25	8.4	0.91*	0.22	0.39	0.81

Cells marked with an asterisk indicate a significant benefit of the site map over the no aids condition ($p < 0.05$) in terms of reduced time, fewer steps and reduced navigational redundancy.

Table 6
Comparison of basic performance measures with and without the bookmark available

Site metaphor	Pacing condition	Differences from no-aids condition				absolute measures	
		*T*1 (s)	*St*1	*Nr*1	*Bm_al*	*Ef*	*Tg*2
Spatial	Unpaced	97	29.1*	0.82*	6.1	0.88*	1.00
	Paced	−8	8.4	0.31*	4.2	0.63	0.90
Non-spatial	Unpaced	−160	3.3	0.05	4.8	0.57	0.89*
	Paced	−154	2.2	0.25	4.1	0.51	0.77

Cells marked with an asterisk indicate a significant deviation ($p<0.05$) in performance compared to the equivalent condition with no navigational aids present. Negative values demonstrate a performance decrement.

The second point of note is that the effectiveness in the use of these aids appears to be related to the general difficulty of the task, as determined by context and time pressure. Table 5 illustrates a range of parameters of performance using the site map in comparison to performance in its absence. The first three: $\Delta T1$ — difference in time of initial search; $\Delta St1$—difference in the number of steps in initial search; and $\Delta Nr1$—difference in navigational redundancy illustrate gains over the no-aids condition. Thus, $\Delta T1 = T1_{(\text{no aids})}$—$T_{1(\text{site map available})}$, for example, and positive values indicate performance gains with tool use. Those marked with an asterisk represent significant levels of gain at the 5% level when compared to the no-aids condition. The remaining three parameters (*M_Pr1* — the proportion of initial search pages using the map; *Ef* — the efficiency of search in the retrieval stage; and *Tg2* — the proportion of targets accessed at retrieval) are absolute measures. *M_Pr*1 cannot be compared to the no-aids condition for obvious reasons, and the other two measures marginally fail to demonstrate a significant advantage of the site map. However, all three measures are consistent with the pattern in which the effects of time pressure and non-spatial task context increase the task difficulty and reduce performance. Of specific interest here is the use of the site map in search (*M_Pr*1) which illustrates that the map is actually used less in the harder conditions, with a concomitant reduction in performance in other measures such as the time taken to complete search, and navigational efficiency in the second phase of the task.

A similar performance profile is seen in the use of the bookmark tool, illustrated in Table 6, where *M_ Pr*1 is replaced with *Bm_al* — the mean number of bookmarks used. The significant observation is that, as in the use of site maps, users apply bookmarks to fewer targets as the task becomes difficult. Note also that, as the tasks become more demanding with pacing and/or a non-spatial metaphor, users actually *slow down* when using bookmarks in comparison to the no-aids condition. Not merely are the advantages of the navigation tool lost in these cases, but its use appears to be a significant burden in itself. To summarize, with either navigation aid, their use represents an immediate commitment of cognitive resources (which in the harder conditions may already be stretched) for longer term gain. We have already

Table 7
Frequencies of questionnaire responses to one of four categories of disorientation in response to the question "how often did you feel lost whilst navigating?"

Questionnaire categories	Unpaced	Paced	Spatial	Non-spatial
Several times"	21	36	18	39
"Sometimes"	45	41	36	50
"Rarely"	26	26	32	20
"Never"	23	12	24	11

Data are split for paced and unpaced conditions and for spatial and non-spatial metaphor conditions.

seen in the above analyses that their use is less likely in harder task conditions and search efficiency in the second phase of the task suffers accordingly.

In these tasks, there is another cognitive load to consider, which is that users have to keep track of task status: where they are, how many targets they have found, how to remember where they are, and in the final retrieval stage how many targets remain to be retrieved. Beyond the performance measures described above, subjective data from the questionnaire indicate that the cognitive activities of maintaining the status of the task were harder under increasing task load. Compared to those using the non-spatial metaphor sites , participants using the spatial metaphor sites were more likely to report having used a strategy[3] to ensure that all five targets in the initial search target were located [95% vs. 80%, χ^2 (1df, $N = 230$) = 13.51 ($p < 0.001$)] and similarly in the retrieval task [97% vs 86%, χ^2 (1df, $N = 230$) = 10.46 ($p < 0.001$)]. The same pattern is repeated in the comparison of non-paced and paced participants, with both the initial unpaced search reporting more strategic behaviour than in the paced conditions [96% vs. 79%, χ^2 (1df, $N = 230$) = 15.34 ($p < 0.001$)] and for the subsequent retrieval task [96% vs 87%, χ^2 (1df, $N = 230$) = 5.62 ($p < 0.025$)].

Participants were also asked to indicate the extent to which they had felt lost whilst navigating. The frequencies of their responses, split both as a function of metaphor and conditions of time pressure, are given in Table 7. The comparison of spatial and non-spatial metaphor shows a significant increase in disorientation in the non-spatial metaphor condition [χ^2 (3df, $N = 230$) = 17.52 ($p < 0.005$)]. Equally, participants are more likely to report disorientation in paced conditions, although this comparison marginally fails to reach statistical significance [χ^2 (3df, $N = 230$) = 7.70 ($p < 0.1$)].

Other questionnaire responses indicate a diminution of cognitive resource available for managing the status of the task, although the numbers involved preclude reliable statistical tests. For example, in the conditions using site maps,

[3] These strategies are varied, but might include counting the number of targets located or rehearsing a string of located targets in sequence, for example.

participants frequently reported attempting to remember the location of found targets on the map; in which case recall would enable the user to use the site map in the same way as a bookmark list in the retrieval stage. In the spatial metaphor, 70% of the users (21 out of 30) report this strategy, compared to 47% (14 out of 30) in the non-spatial metaphor condition. In non-paced conditions, this strategy is reported by 19 participants compared to 16 in the paced condition. Neither comparison reaches statistical significance. Nevertheless, the overall picture is consistent with the view that effective use of tools, cognitive maintenance of task status, and planning for the retrieval stage (particularly in the use of bookmarks) are all likely to be impaired by task stressors.

The final issue to consider is that of what happens when both navigation aids are available. In this, the performance in the non-spatial context is not difficult to explain, since the effect of the two aids appears to be additive: the map appears to increase the navigational efficiency of the task and the bookmarks (albeit fewer than the ideal) give access to more targets quickly. In the spatial context case, on the other hand, the bookmark condition gives higher efficiency of search and a faster initial search than do both aids together. Compared to the map condition, the both aids condition is slower and offers no gain in search efficiency. This is illustrated in Fig. 1 and the previous analyses reported. In the spatial context condition, the effect of having two aids is subtractive rather than additive.

We offer an explanation based largely upon the participants' questionnaire responses and a small measure of special pleading based upon a recurrent observation by several participants. First, Table 8 summarizes their responses to the question of how they attempted to ensure that all five targets had been found in the first part of the task as a function of which combinations of aids were available; other factors being collapsed for this analysis. Note that the users with bookmarks available are predominately using this as an external memory — targets are found and bookmarked and the list acts as an easily available record. Without this, users typically count or rehearse the target screens as found. However, it is noteworthy that in the condition with both aids available this is not done; despite this being the predominant strategy associated with the map. Table 9 collates participants' reported strategies for ensuring that they will remember the targets for the second stage of the task. Here, those with site maps alone attempt to use the location in the

Table 8
A collation of users' strategies for ensuring that they had found all the principal targets in the first phase (search) of the task as a function of the navigation tools available to them

Description of strategy	No aids	Site map	Bookmark	Both aids
Bookmark target screens and count the number of bookmarks used	n/a	n/a	47	46
Rehearse target names	20	5	0	0
Count target screens as they are found	23	46	7	7
No strategy reported	7	9	7	7

Table 9
A collation of users' strategies for ensuring that they remember the location of targets in the second phase (retrieval) of the task as a function of the navigation tools available to them

Description of strategy	No aids	Site map	Bookmark	Both aids
Memorize the location of targets on the site map	n/a	35	n/a	0
Recognize the target names when found	25	14	5	7
Rehearse target names during retrieval	11	3	4	2
No strategy reported	14	8	51	51

site map as a memory cue. Those with a bookmark alone need no strategy since the bookmark list contains all necessary information. However, in the condition using both aids, participants make no attempt to use locations in the site map as a memory cue; relying wholly on the bookmark list. The problem in this is that users in the "both aids" condition use significantly fewer bookmarks than do users in the bookmark alone condition [an average of 2.9 compared to 4.9 F (1,118) = 26.03 ($p < 0.0001$)]. This effect is particularly marked in the non-time-pressured, spatial context condition (i.e. the easiest), where use of the bookmarks alone is practically optimal when only the bookmark is available and is reduced by half when both navigation tools are available.

To summarize, users with both navigation aids in the spatial context condition fail to use either efficiently. They neither use the map as a memory cue, and nor do they use the bookmarks as comprehensively as in the condition when only the bookmarks are available. Consequently, to access those targets with no bookmarks, they lack any information to focus their search, leading to an inefficient second search.

Why do users do this and why not in the non-spatial condition? Participants comments strongly suggest that the issue rests upon the users perceived equation of costs in the following way. With a bookmark as the only navigation tool, navigation is generally difficult and planning for the retrieval process leads users to bookmark targets effectively during the original search. Although bookmarking takes time and thought, this is perceived correctly as worthwhile. In conditions where only the site map is available, although navigation is relatively easy, it appears that users recognize the value of remembering map locations as substitute bookmarks. However, when both are placed together, the users do not use the map as a memory aid and make fewer bookmarks. Together, the combined facility of the tools means the user exploits neither efficiently and their combined effect is subtractive, not additive.

4. Discussion

This study is both an analysis of the usability of navigating tools and a psychological analysis of how these tools are used as a function of circumstance. Considering the usability perspective first, an objective of this study was to compare and contrast the use of bookmarks with site maps. This was partly motivated by the observation that bookmarking and site maps as navigating tools offer different opportunities for design and customization in interfaces. Bookmarking tools are usually available as part of an information environment, and as such are independent of its information content. Site maps, on the other hand, are more context dependent. In heterogeneous information structures such as the World Wide Web, site maps, when used, are designed by individual providers to service their site only. This raises a set of design issues about where site maps are effective; how sites might be designed to best exploit them and how they interact in use with other navigation facilities. We consider these below in the light of the results we have reported.

It is a necessary, if fairly obvious, preliminary to emphasize that the *utility* of these tools in themselves is not in question. This study, like others before, confirms that bookmarks enable direct access of bookmarked targets and site maps enhance the ease of none-to-node navigation. Thus users are quicker during search with site maps and are less likely to report being lost or disorientated. On the other hand, they are more efficient during retrieval using bookmarks even though this tool has no impact upon the users' subjective sense of orientation. Any other findings would perhaps have been surprising, since this is what these tools are designed to do. Note, however, that in isolation — and more markedly in the less demanding conditions — both navigation tools are capable of being used adaptively by the users. Thus bookmarks can reduce search time because they can be used to jump around the site, and site maps are used like bookmarks when users remember the locations of targets on the map and access them directly.

This ability to adapt the use of navigation tools is strongly linked to the context and is most clearly seen in the spatial task context in the absence of time pressure. In these conditions, users appear to have sufficient cognitive resources to act strategically upon the task environment in order to use bookmarks effectively both as tags for retrieval and for navigation aids and to use site maps as navigation aids and *aides memoire* for the retrieval task. In other conditions, this is less apparent, although not absent. At this point, however, we encounter an irony in user performance. This adaption is observed when either navigation aid is used singly. When present in the same interface, users appear to be less efficient in using either the bookmarks or the site map, with the result that initial search is slower and retrieval performance in the this condition is poorer than if users had the bookmark alone.

Why should users of two navigation tools be inefficient planners in the spatial task context? In this respect, users' subjective judgements of orientation are revealing. As noted above, site maps produce a significant improvement in this respect whereas bookmarks, unsurprisingly, have no impact. However, Tables 10a and b show the outcome of post hoc tests comparing subjective estimates of disorientation in the different conditions of tool availability. Note, as already stated, that this measure does not vary significantly between the no-aids condition and the bookmark condition; i.e. the bookmark does not lead users to feel any less disoriented. However, comparing users of the site map to those where both navigation aids are

Table 10
(a) Frequencies of questionnaire responses to one of four categories of disorientation in response to the question "how often did you feel lost whilst navigating?" as a function of the navigation aids available to the user

Questionnairre categories	No aids	Site map	Bookmark	Both aids
"Several times"	23	10	21	3
"Sometimes"	20	21	27	18
"Rarely"	4	12	9	27
"Never"	3	17	3	12

(b) Chi-square comparisons of the frequencies of questionnaire categories for the different combinations of navigation aids. Each test has 3df (5% crit = 7.82) and only the comparison of no aids and bookmark fails to reach significance

	No aids	Site map	Bookmark	Both aids
No aids	—	19.38	2.20	41.56
Site map		—	15.99	10.99
Bookmark			—	32.20
Both aids				—

present, they are even *less* likely to report feeling lost or disoriented with both aids compared to the site map alone. It is as if bookmarks — apparently with little or no benefit to orientation on their own — give an illusion of added value when accompanied by a site map; with the result that users are lulled into planning less. In what is essentially a cost–benefit analysis, users are failing to apprehend the benefits of planning (in this case using bookmarks and remembering site map locations — each with their attendant costs) in the face of the apparently low costs of the current search.

Considering the implications of our findings for design, we must first make a distinction between findings which imply useful and general design heuristics and those which warn of task and context specificity — potential difficulties in design which can only be anticipated in a general sense and not necessarily prescribed in the particular case. On the face of it, the demonstrated benefits of spatial metaphors for site structure reinforces previous claims (e.g. Hammond and Allinson, 1987; Smilowitz, 1996; Kim and Hirtle, 1995; Kim, 1999) for this as a design heuristic. On the other hand, the complexities of tool use as a function of context are a good example of context specificity.

Dealing first with the issue of context specificity, it is, of course, frustrating for designers, in response to the question: "What is the best navigation tool" to be told that it depends on the circumstances. However, in this study we have attempted to analyse something of the complex relationship between task, context and navigation tools. This points to two answering two questions to inform difficult design decisions that will hinge upon the specifics of a particular application.

(1) *What is navigation for?* This study establishes the need to distinguish between search processes and retrieval processes when designing tools for specific tasks. Activites which concentrate upon searching for targets on a one-off basis place a design premium upon the facility of movement and are less concerned with subsequent recovery. In such cases, learning about the structure of sites and planning activites for subsequent retrieval are of lower value and there is less need to design with these objectives in mind. A converse argument applies to applications in which search is a relatively rare activity and repeated recovery of known targets is at a premium. If we accept Tauscher and Greenberg's (1997) findings, this is hardly uncommon.

(2) *Do we need the users to plan?* It appears that if we do not consider the users' planning behaviour, the provision of two navigation tools can sometimes be subtractive, not additive, to performance. In the non-spatial task conditions, the design premium was upon ameliorating as many of the task difficulties as possible to reduce the cognitive overload of the task, and, in these experiments, the effect of having two tools available was additive. Using the spatial task metaphor, on the other hand, the two tools are not additive; and it might even be desirable in some circumstances to make the task *harder* in order to improve overall performance in the task. The critical issue here is some tasks, which are nominally navigation (such as getting to a specific site or page) benefit from planning or other activities in which a preliminary investment of time and effort returns a performance profit. Apart from a general reluctance to trade-off immediate effort against future gain [see also Malone (1983) and Lansdale and Edmonds (1992) considering information management strategies], this study identifies *two* further factors which influence this issue. First, users must have sufficient cognitive resources free to make appropriate planning behaviour possible. The theory of interface design is becoming reasonably well-equipped to help in this respect. Second, even if those resources are available, and whatever the apparent ease of a task, the users need to be able to recognize when extra investment is worthwhile in the face of apparent ease of use. Research in this area of interface design is still in its infancy. The best current research can do at present is to warn of this complex interaction of performance factors and allow designers to consider these issues as they arise.

Returning to the specific findings about navigation tool use, our previous analyses have distinguished between search and retrieval processes in the two phases of our experiments and we have used, amongst others, the completion times for these phases (T_1 and T_2) as measures of performance. However, the users' task was to complete both phases. Therefore, the composite measure ($T_1 + T_2$) might be a more useful index of performance by which to compare conditions overall.[4] Fig. 2 presents a scatterplot for the eight conditions of tool use (none, site maps, bookmarks or both aids in both spatial and non-spatial metaphors) with the abscissa representing ($T_1 + T_2$) for the unpaced conditions and the ordinate for the paced conditions.

The first point of note is the close correspondence of all conditions to the theoretical function of parity between paced and unpaced conditions. In effect, this

[4] We are grateful to Stephen Payne for this observation.

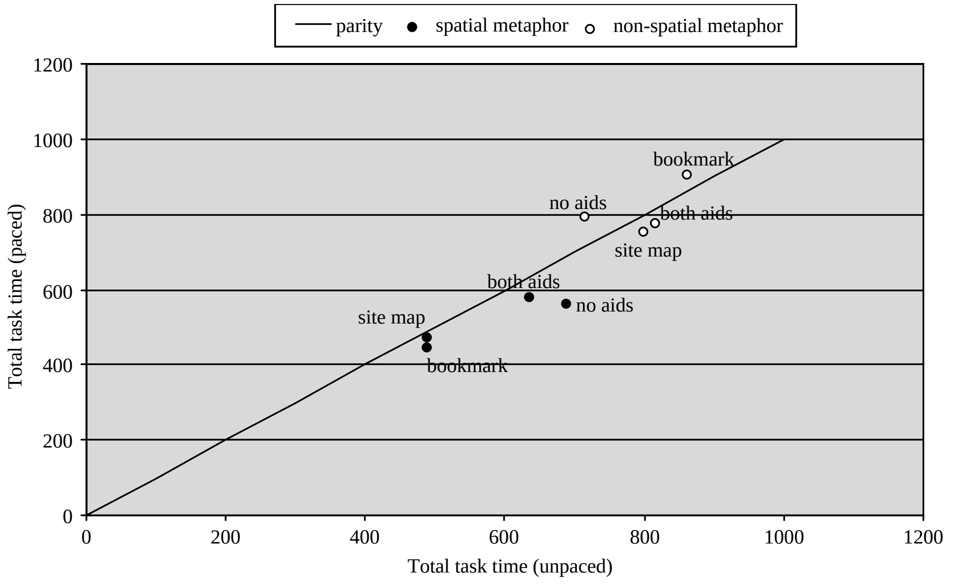

Fig. 2. A scatterplot relating total task time (T1 + T2) in paced and unpaced conditions of tool availability and site

suggests that the time saved in the search phase of paced conditions is exactly balanced by the extra time required to carry out the retrieval stage. We see no *principled* reason why this should be, or whether it will apply to other circumstances. There seems no good reason why, for example, the exact time spent bookmarking a target should be matched by the extra searching in the second phase of the task if that bookmark was not made. Nevertheless, this outcome is useful for the present purposes because it allows us to rank-order these conditions and tool combinations in terms of this performance measure. Note also that the variations between conditions are large and not second-order effects. For example, total task completion time in the spatial metaphor condition with both navigation aids available is approximately 150 s greater than that for either tool alone, representing , an increase of some 33%.

Considering again the relative usability of these two navigation aids, this rank-ordering suggests that site maps alone are a "safer" bet in design, given the context-specificity of the bookmarks and the combination of both. With the spatial metaphor, this ranking from most to least efficient is *Bookmark-Map-Both-None* (with the distinction between the first two being negligible), whereas with the non-spatial metaphor the same ranking is *None-Map-Both-Bookmark* (although we should note in the latter case that lower target retrieval rates, particularly in the *None* condition, make these orderings partial). It appears reasonable to conclude that, in the absence of specific intelligence about task-specific factors such as pacing or the representational metaphor, a site map is the single tool most likely to optimize the total time spent searching, and retrieving from, complex information sites. In both the spatial and non-spatial metaphors tested here, this tool produces overall task performance times reliably close to the minimum observed, whereas the bookmark (either alone or combined with a map) produces high performance in some conditions and lower performance in others.

We return now to the design heuristic that sites should be designed according to a spatial metaphor. Notwithstanding the complexities of the observed interactions between the combinations of navigation aids, our experiments indicate that the impact of familiar spatial context in these experiments is very large and gives strong empirical support to this recommendation. As an heuristic, therefore, we feel confident that an information site based upon a spatial metaphor is likely to be navigated more efficiently than an equivalent site based upon a non-spatial metaphor. There is, however, much more research required to be confident about what aspects of the spatial metaphors are delivering this added value. As in the broader area of cognitive science, the reasons why spatial contexts can produce higher performance can be extremely complex and may not necessarily imply that spatiality itself is the critical factor. In this specific case, a number of directions for future research present themselves.

(1) *What is the value of prior knowledge in schematic knowledge?* Although the two web site contexts we used here were identical in structural terms, the spatial context taps into a well-understood schema of house layout. Thus users come with appropriate prior knowledge that stairs lead to bedrooms and garden sheds are unlikely to be found on the landing. Put another way, the coherence of the relationships between different pages in this site was already well established. No such expectations could have applied to whom *Peter* or *Clare* knew in the non-spatial metaphor conditions, and as a result there is an additional implied cognitive load in the non-spatial context. Therefore we cannot say from these results what would happen if a novel spatial schema (for example, streets in a fictional town) was substituted for the house schema. This is important because, in recommending spatial metaphors as design heuristics we need to know whether it is the spatiality or the prior knowledge which determines the effectiveness of the metaphor and this choice in turn determines whether we need to use familiar spatial metaphors for this purpose, which will limit how the site is actually constructed or whether the spatial metaphor will tolerate customization to correspond to the pre-existing structures of the site.

(2) *Is the imagability of the context important?* One benefit of the spatial metaphor may the imagability of the individual rooms. If we think of the search and retrieval task as one in which pages are associated in memory with targets, title names and locations on a site map, it is a well-known finding that the more concrete and imagable the to-be-associated concepts are, the easier the association (Paivio, 1969). Thus a conservatory offers many opportunities to provide concrete elaborations (usually but not necessarily imagistic) of plants, cane furniture, and tiled floors which the name *Clare* might not. Further research may establish whether there is a design premium for making web sites and their individual pages memorable in this way as a means of facilitating users memory for them and the pages/sites to which they are linked.

(3) *Does the link between task and context matter?* In these trials, the task, targets and context are coherent concepts within the same scenario. In the spatial metaphor in particular, the targets (such as a tap) have significance within the context, where some rooms can be expected to have taps and others not. Users are looking for flowers and taps within in a house, or are "talking" to students in a social structure. What happens if that link is violated? For example, what happens if we use a house as a spatial context for a chronological sequence of patient records in a hospital? It is not obvious that useful incidental associations between information and context would necessarily ensue if the scenario of the spatial metaphor was perceived by the user as being meaningless (e.g. see Lansdale, 1991). This might mean that users might find it easy to navigate the site but not as easy to remember what was where.

(4) *How much does spatial coherence matter?* The experimental tasks in these experiments are coherent in that the information structures map exactly onto the task scenario. This will not necessarily be the case in the real world. If answers to the above questions confirm that spatial metaphors are useful in design, we will still find that most information sites map poorly onto them. For example, in most WWW sites it is not possible to represent the transivity of links as nodes in three-dimensional space. In this case, the design decisions are either to decrease the coherence of the metaphor by allowing inconsistencies, to redesign the information structures to conform to a coherent metaphor, possibly with the loss of useful links or some combination of both. More research is required to inform whether, and how, we might make such compromises in design.

Not until these and other questions are answered will we be able to address the question whether it is spatiality, or a byproduct of spatiality, that is the important element in the advantage of spatial metaphors for navigation. It may be that, in the experiments carried out to date, spatiality has been applied in such a way as to increase imageability, the incorporation of prior knowledge, the coherence of contexts and the likelihood that the task and the scenario will be integrated. Thus, while we are happy to speculate that spatial contexts are *more likely* to produce easier sites to navigate and within which to find known targets, we remain unsure exactly why that should be.

Acknowledgements

Stephania Padovani was supported by a doctoral grant from CAPES—Brazilian Federal Agency for Postgraduate Education.

Appendix A.

Definitions of performance measures used are given in Table 11.

Table 11
Definitions of performance measures

Abbreviation	Task phase?	Description/definition
$T1$	1: Search	Total elapsed time (s) to complete the first phase of the task
$T2$	2: Retrieval	Total elapsed time (s) to complete the second phase of the task
$St1$	1: Search	Total number of screens visited on the first phase (including repetitions)
$St2$	2: Retrieval	Total number of screens visited on the second phase (including repetitions)
$M1$	1: Search	Total number of screens visited on the first phase (excluding repetitions)
$M2$	2: Retrieval	Total number of screens visited on the second phase (excluding repetitions)
$Nr1$	1: Search	Navigation redundancy on the first phase of the task (from Gomes, 1994) $\mathrm{Nr1} = (St1—M1)/St1$
$Bk1$	1: Search	Proportion of Screens accessed by the "back" button $Bk1 =$ (No. of uses of 'back button)$/St1$
$Tg1$	1: Search	Proportion of targets acquired during the search task $Tg1 =$ (no of targets found)/maximum number
$Tg2$	2: Retrieval	Proportion of targets acquired during the retrieval task $Tg2 =$ (no of targets found)/maximum number
Ef	2: Retrieval	Navigational efficiency on the second phase of the task $Ef = Tg2$ (no. of screens on optimal retrieval route)$/St2$
Usr	2: Retrieval	Unnecessary screens visited on the second (retrieval) stage of the task $Usr =$ (no. of screens on optimal retrieval route)$/St2$
M_Pr1	1: Search	Proportion of screens accessed via the site map on the search phase $M\_Pr1 =$ (no. of uses of site map)$/St1$
Bm_a1	2: Retrieval	Proportion of screens accessed via the bookmark on the search phase $Bm\_a1 =$ (no. of uses of Bookmark)$/St1$

References

Baddeley, A.D., 1963. A zeigarnik-like effect in the recall of anagram solutions. Quarterly Journal of Experimental Psychology 15, 63–64.

Bower, G.H., Clark, M.C., Lesgold, A.M., Winzenz, D., 1969. Hierarchical retrieval schemes in recall of categorized word lists. Journal of Verbal Learning and Verbal Behaviour 8, 823–843.

Chase W.G., Ericsson, K.A., 1981. Skilled memory. In Anderson, J.R. (Ed.), Cognitive Skills and their Acquisition. Erlbaum, Hillsdale, NJ.

Chen, C., Rada, R., 1996. Interacting with hypertext: a meta-analysis of experimental studies. Human-Computer Interaction 11, 125–156.

Chou, C., Lin, H., 1998. The effect of navigation map types and cognitive styles on learners' performance in a computer-networked hypertext learning system. Journal of Educational Multimedia and Hypermedia 7, 151–176.

Conklin, J., 1987. Hypertext: an introduction and survey. IEEE Computer 20, 17–41.

Darken, R.P., Sibert, J.L., 1996. Wayfinding strategies and behaviours in large virtual worlds. In: Proceedings of the CHI'96 Conference, http://www.acm.org/ sigchi/chi96/proceedings/papers.html.

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

Godden, D.R., Baddeley, A.D., 1975. Context-dependent memory in two natural environments: on land and underwater. British Journal of Psychology 66, 325–331.

Gomes, M.J., 1994. Um estudo sobre comportamento de navegação em hiperdocumentos - uma abordagem quantitativa. M.Sc. Thesis, Universidade do Minho, Portugal.

Hammond, N., Allinson, L., 1987. The travel metaphor as design principle and training for navigating around complex systems. In: People and Computers, III: Proceedings of HCI '87 Conference, (pp. 75–90). Cambridge: Cambridge University Press.

Hammond, N., Allinson, L., 1989. Extending hypertext for learning: an investigation of access and guidance tools. In: People and Computers V: Proceedings of HCI '89 Conference. Cambridge University Press, Cambridge, pp. 293–304.

Hyde, T.S., Jenkins, J.J., 1973. Recall for words as a function of semantic, graphic and syntactic orienting tasks. Journal of Verbal Learning and Verbal Behaviour 12, 471–480.

Kim, J., 1999. An empirical study of navigation aids in customer interfaces. Behaviour & Information Technology 18 (3), 213–224.

Kim, H., Hirtle, S.C., 1995. Spatial metaphors and disorientation in hypertext browsing. Behaviour and Information Technology 14, 239–250.

Lane, D.M., Robertson, L., 1979. The generality of the levels of processing hypothesis: an application to memory for chess positions. Memory and Cognition 7, 253–256.

Lansdale, M.W., 1991. Remembering about documents: memory for appearance, format, and location. Ergonomics 34, 1161–1178.

Lansdale, M.W., Edmonds, E.A., 1992. Using memory for events in the design of personal filing systems. International Journal of Man-Machine Studies 36, 97–126.

Malone, T.W., 1983. How do people organise their desks? implications for the design of office information systems. ACM Transactions on Office Systems 11, 99–112.

Marchionini, G., Shneiderman, B., 1988. Finding facts versus browsing knowledge in hypertext systems. IEEE Computer 21, 70–81.

McAleese, R., 1989. Navigation and browsing in hypertext. In: McAleese, R. (Ed.), Hypertext: Theory into Practice. Intellect Books, Oxford, pp. 5–38.

McDonald, S., Stevenson, R.J., 1998. Effects of text structure and prior knowledge of the learner on navigation in hypertext. Human Factors 40, 18–27.

Miller, G.A., 1956. The magical number seven plus or minus two: some limits on our capacity for processing information. Psychological Review 63, 81–97.

Morris, C.D., Bransford, J.D., Franks, J.J., 1977. Levels of processing versus transfer appropriate processing. Journal of Verbal Learning and Verbal Behaviour 16, 519–533.

Nielsen, J. (1995). Navigating Large Information Spaces. In: Nielsen, J. (Ed.) Multimedia and Hypertext: the Internet and Beyond. London, AP Professional, pp. 247-272.

O'Hara, K.P., Payne, S.J., 1998. The effects of operator implementation cost on planfulness of problem solving and learning. Cognitive Psychology 35, 34–70.

Padovani, S., 2001. The effect of navigational aids on users' navigation strategies. Unpublished Ph.D. Thesis, Loughborough University, UK.

Paivio, A., 1969. Mental imagery in associative learning and memory. Psychological Review 76, 241–263.

Paolucci, R., 1998. The effect of cognitive style and knowledge structure on performance using a hypermedia learning system. Journal of Educational Multimedia and Hypermedia 7, 123–150.

Smilowitz, E.D., 1996. Do metaphors make Web browsers easier to use? In Designing for the Web: Empirical Studies (2nd Microsoft Conference in Human Factors and the Web), http://www.baddesigns.com/mswebcnf.htm.

Tauscher, L., Greenberg, S., 1997. How people revisit web pages: empirical findings and implications for the design of history systems. International Journal of Human-Computer Studies 47, 97–137.

Wright, P., Lickorish, A., 1990. An empirical comparison of two navigation systems for two hypertexts. In: McAleese, R., Green, C. (Eds.), Hypertext: State of the Art. Intellect Limited, Oxford, pp. 84–93.