Does Location Come for Free? The Effects of Navigation Aids on Location Learning

Carl Gutwin and Diana Anton
Department of Computer Science, University of Saskatchewan
gutwin@cs.usask.ca
HCI-TR-06-03

ABSTRACT
Navigation aids such as bookmarks, target prediction, or history mechanisms help users find desired objects in visual workspaces. They work by highlighting objects that may be important, and they can improve performance in spaces where the territory is not well known. However, by making navigation easier, they may also hinder acquisition of a mental map of the space, reducing navigation performance when the navigation aid is not available. We carried out a study to determine the effects of three different types of navigation aids on spatial location learning. We found that after training with a navigation aid, there was ***no*** reduction in performance when the aid was removed. Even with training interfaces that made the task significantly easier, people learned the locations as well as those who had no aid at all in training. These results suggest that designers can use navigation aids to assist inexperienced users, without compromising the eventual acquisition of a spatial map.

Author Keywords
Location learning, spatial memory, navigation aids, revisitation, prediction, bookmarks, visit wear, read wear.

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

Many visual applications and visualization systems present graphical workspaces with many different objects that can be visited and inspected by the user. In these systems, there is often not enough screen space to adequately label each item, and so users cannot reliably look for things using visual search; instead, they must find things by a combination of inspection and memory (see Figure 1).

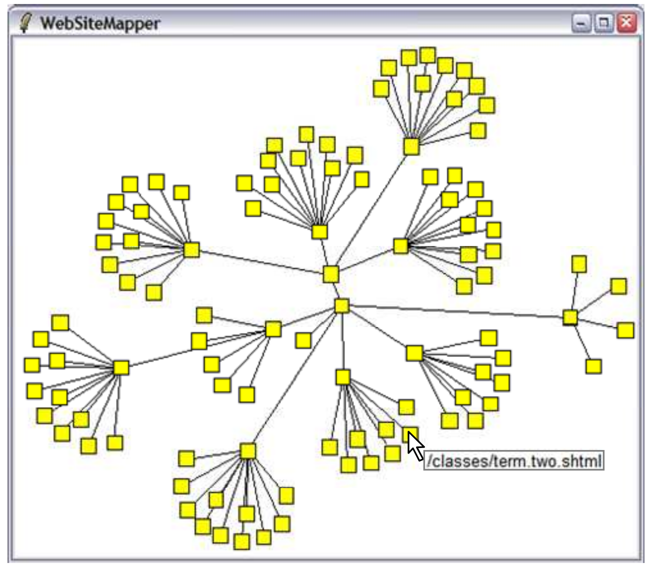

Figure 1. A web-site visualization system where there is not enough room to put page labels. A popup tag follows the cursor to allow inspection of nodes.

Inspection is slow, however, and when users are not familiar with the dataset, it can be difficult for them to find the objects that they need. To help with this problem, several systems provide *navigation aids* – visual markers to highlight objects that may be important. Marking objects can be based on a variety of information depending on the domain: for example, read wear to show the items that the user has previously visited, flags or bookmarks that allow the user to explicitly mark items, or predictive interfaces that suggest which items are likely to be used next.

Navigation aids can dramatically improve performance if they accurately indicate the objects that really are important for the user's task. When successful, they allow the user to carry out a much simpler task – use visual search to find those objects that are highlighted, and then use inspection or memory to find the desired object within that much smaller set. In the limit, a navigation aid would always highlight the exact object that the user needed next, reducing the problem to a simple perceptual search (e.g., 'always click on the red object').

The problem with navigation aids is that they are not always correct, and not always available. This can happen for several reasons: for example, a prediction interface might simply get the prediction wrong; the user may switch to a different application that does not display their flags and annotations, or a read-wear effect may have faded by the time the user needs to return to the object.

What happens to performance when a user learns to depend on the visual aid, and then has to work without it? Navigation aids make the task easier in the short term, but does this reduction in effort mean that people will fail to build a long-term mental map of the territory? Previous research by Ehret [7] suggests that if a user's retrieval strategy does not explicitly involve location knowledge, they will learn the object locations much more slowly. That is, a perfect 'follow the red highlight' strategy would result in the least location learning.

This presents a dilemma for designers: navigation aids are valuable when users are unfamiliar with a dataset; but if the eventual goal is for users to memorize locations and create a spatial map, it is possible that navigation aids are a detriment.

To determine whether navigation aids really do hamper spatial location learning, we carried out an experiment in which people were asked to repeatedly find items on the screen. People trained using one of four retrieval interfaces: one that showed the object's labels, a predictive interface that was 50% accurate, a predictor that was 90% accurate, and an interface with no navigation aid at all.

We found that during training, the navigation aids were effective: both the interface with object labels and the interface with 90% prediction significantly improved performance compared to retrieving objects with location only. Once the navigation aid was turned off, however, we found **no** difference in testing times between those who had practiced with a navigation aid and those who had practiced with locations only. Even when people had trained with a 90%-accurate predictor, they were as fast at retrieving objects using location alone, as those who had trained with locations from the start. A power analysis carried out before the study provides evidence that the lack of difference is reliable.

BACKGROUND

There are two main areas of related work that are relevant to our investigation: navigation aids, and spatial object location memory.

Navigation Aids

Navigation aids in graphical workspaces are visual markers that are intended to assist the user in determining where they have been, where they are, or where they want to go next. Navigation aids work in two ways: by reducing the number of items that must be considered in a visual or inspection-based search, and by providing additional landmarks in the space that assist spatial memory (e.g., "it was just past the red flag").

There are several possible types of navigation aid in visual workspaces, including trails, edit wear, flags, bookmarks, and prediction systems. We divide these into two main groups: predictive markers and user-chosen markers.

Automatic / Predictive markers

Some navigation aids are based on models or predictions about what objects the user is going to want to visit. The predictions can come from a variety of sources:

- pre-selected sets of important objects (e.g., the geographical locations shown in Figure 3),
- markers arising from previous user actions (e.g., halos of off-screen objects from a previous query [3]),
- direction arrows in games (since the game knows where the goal is located),
- markers based on interaction history, such as read wear [10] to help people return to previously-visited objects (see Figure 2), or trails built up from others' activity [22],
- artificial landmarks to provide additional visual distinctiveness to objects [13], or
- prediction based on task knowledge such as dependencies between items.

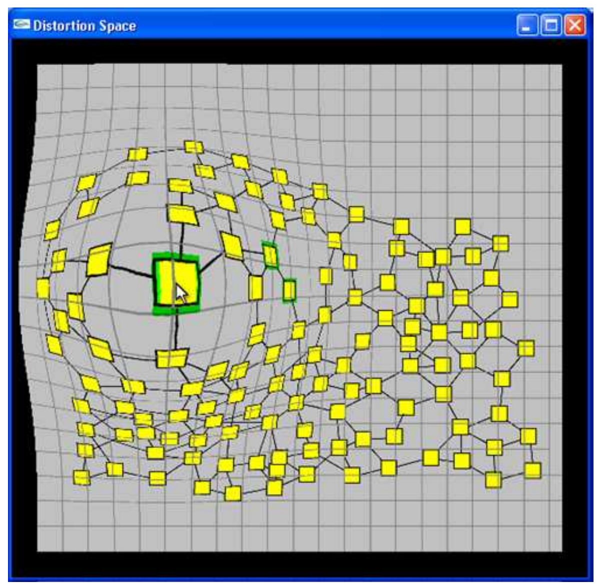

Figure 2. Example navigation aid in a fisheye system, based on interaction history (from [19]): the last three nodes visited by the user are highlighted with a green border.

User-chosen markers

Other navigation aids are chosen and placed by the user themselves. These are common in the real world (flags, 'breadcrumb trails,' blazes), and appear in many computational spaces. Examples include 'brushing' in scatterplot visualization systems [4], where objects are coloured so that they can be found again when the view changes, user-chosen flags in mapping systems (e.g., Figure 3), and paths intentionally created through a set of web pages [22].

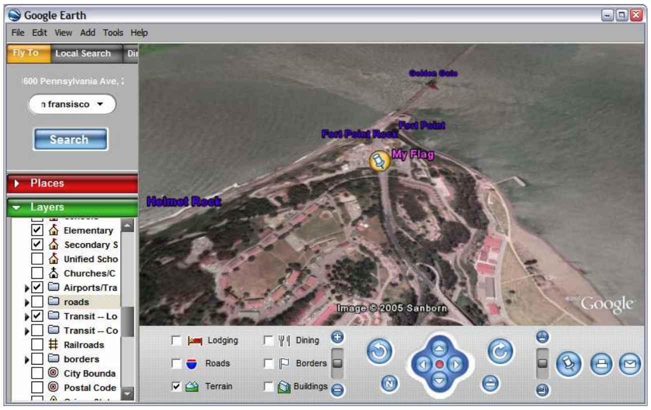

Figure 3. A mapping application (www.google.com) with pre-selected markers (in blue) and a user-selected flag (centre of picture).

The potential problem with navigation aids, however, is that they may provide an easier navigation strategy than that of remembering objects' locations – meaning that people may start to depend on them for navigation. When the aid is not present (e.g., the prediction is wrong, the user defined flags are turned off, or the user is working on a different machine), people may be unable to navigate effectively. An analogue for this problem involves people's experiences with 'speed dial' on telephones – that is, if a person's number is in the speed-dial list, the user never learns the actual telephone number. On a phone without the speed-dial mapping, the user is unable to complete the task.

Spatial Memory in Computer Interfaces

Several researchers in HCI have explored the use of spatial memory in computer interfaces (e.g., [6,7,12,16]), and studies have shown that although abilities can vary widely [18], people are capable of using spatial memory to remember large numbers of items.

For example, Robertson and colleagues [16] tested a spatial memory technique (the Data Mountain) against several other methods of recalling web pages. Retrieval of 100 web pages was significantly faster with the spatial technique than with a standard bookmarking system. The spatial memory also persisted over a long time: participants who returned for a follow-up study several months later were able to retrieve items at the same performance level, with only a brief retraining period [5].

However, other research suggests that when location is used as the only retrieval strategy, spatial memory fares less well. An early study by Jones and Dumais [12] showed that retrieval of items using location alone was slower and less accurate than when items were also represented by name.

Spatial Object Location Memory

A great deal of research has been done on the capabilities and limitations of human spatial memory, as well as on its underlying mechanisms and computations (e.g., see survey in [11]). There are several types of spatial cognition and memory that have been identified, such as the visuo-spatial sketchpad which encodes short-term spatial information (e.g., [2,15]), memory in wayfinding and navigation (e.g., [6,21]), and memory of *object locations* (e.g., [9,15]), which is the area most relevant to the current investigation.

Location memory is memory for the positions of objects in a space. Two different types of location memory have been proposed [11]: first, relative-position memory, where people remember locations in relation to other objects or landmarks in the scene; and second, absolute-position memory, where people remember objects' exact locations (for example, such that they could be found in the dark).

A main issue in spatial memory research is the way in which people learn and retain memory of object locations. Two different possibilities have been explored – that location learning is incidental and almost automatic, and that location learning improves with attention and effort.

Incidental Location Learning

Several studies suggest that at least some aspects of location learning occur automatically, while people carry out other tasks. For example, Andrade and Meudell showed that recall of word *locations* was unaffected by the difficulty of a concurrent task, even though word *recognition* was affected [1]. Another example involves the Data Mountain: Czerwinski and colleagues showed that once participants had learned to retrieve web pages by name and thumbnail, retrieval without either of these cues was only briefly disrupted [5].

Intention and Effort in Location Learning

Other studies provide evidence for the importance of intention in the learning of object locations. For example, a study by Van Asselen et al. demonstrated that when people focused their attention on a route through a building, they were better able to draw a map of the path [21].

A second example is Ehret's study of the effects of visual representations on spatial memory [7], and since this study is used below, we provide additional details here. Ehret agrees that locations are learned incidentally: "without specific intent to do so, users can gradually learn the locations of the interface objects to which they attend" (p. 211). However, he states that *unintentional* is not the same thing as *effortless*, and that the way in which a user interacts with the data has a substantial effect on their resulting location learning.

His studies ask participants to find and select screen objects, where the objects are represented in different ways that require different amounts of location knowledge to perform the task. For example, participants were shown a colour, and had to find a corresponding object on the screen; in one condition, the objects themselves were coloured, and in another, objects were unmarked (but could be inspected by holding the mouse cursor over them). In this case, the colour-matching representation requires very little location knowledge when repeating the task, and the unmarked objects require considerably more.

Ehret suggests that since explicitly remembering locations requires significant effort, people will choose a lower-cost strategy when possible, and their location learning will be impaired as a result. The study showed that people who used the colour-match condition made significantly more errors in reconstructing the map of objects on the screen (however, no differences were found between conditions that used textual labels, icons, or no markings at all).

Although Ehret's work suggests that there are no shortcuts to learning locations, his results are not conclusive for deciding on the effects of navigation aids. In order to further explore this issue, we carried out a study that was based on Ehret's methods and tasks.

STUDY METHODOLOGY

Participants

Sixteen participants (8 men and 8 women) were recruited from a local university. Participants ranged in age from 19 to 31 years (mean of 23 years). All were familiar with mouse-and-windows applications (average use of more than 8 hours per week).

Apparatus

A custom study system was built using Tcl/Tk (see Figures 4-6). The system presented ten rectangular targets arranged in a ring; a box at the centre of the ring was used to display the next object name for retrieval. Users could click on any of the targets to select them. If they held their mouse cursor over the target for one second, a popup tag with the target's name appeared under the item (see video figure). When the user selected the correct target, the target would turn green; incorrect selections caused the target to turn red. The study was conducted on a P4 Windows system with a standard optical mouse and a 1024x768 display.

Experimental Conditions

Four retrieval interfaces were used in the study – three that provided a navigation aid, and one that provided no assistance. The four interfaces require support different retrieval strategies that require varying degrees of location knowledge. The only strategy for finding objects that worked in all conditions was inspection – holding the mouse over an object for one second to see its name. The interfaces were:

- *Labels*. Items on the screen were annotated with permanent labels showing their names (see Figure 4). Therefore, participants could retrieve objects either by remembering locations or by looking for the label. This condition required the smallest amount of location knowledge (since the labels could always be used).
- *Predict-90*. In each trial, a simulated prediction system highlighted one item on the screen with a 5-pixel-wide blue border (see Figure 5). The accuracy of the prediction system could be controlled programmatically (since all targets were known in advance); the Predict-90 condition was 90% accurate (note that this represents average accuracy; within any one block, prediction could be higher or lower). Participants could find an object by inspecting the highlighted item or by remembering locations. Since following the highlight was correct 9 out of 10 times, participants in this condition did not require location knowledge in order to be efficient.
- *Predict-50*. The Predict-50 condition was identical to the Predict-90 system, except that the predictor was only 50% accurate (average accuracy). This condition required location knowledge in half the trials.
- *Location-only*. This condition provided no visual aid at all; items on the screen were not marked in any way (see Figure 6). The only way to find an object in this condition was either to remember the location, or inspect the objects.

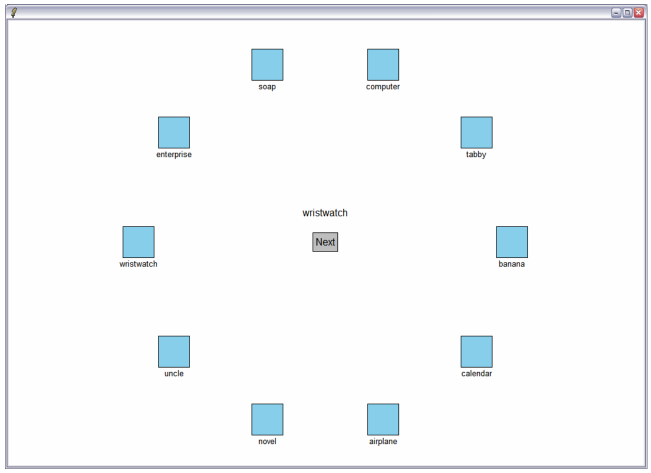

Figure 4. Experiment system in Labels condition. Item to be found and selected next is given by name in the centre.

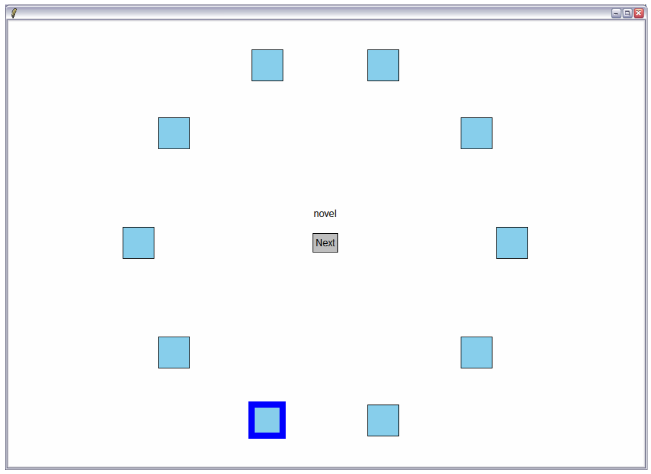

Figure 5. Experiment system in Predict conditions (accuracy of the highlight was either 50% or 90%).

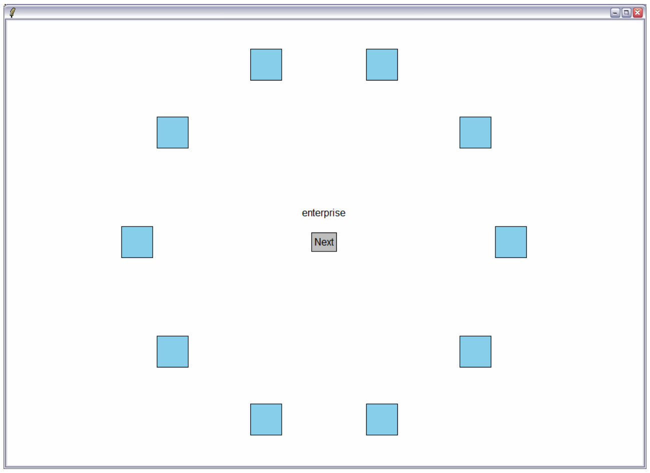

Figure 6. Experiment system in Location-only condition

Tasks and Datasets

The items on the screen were given arbitrary names taken from lists of everyday objects (e.g., camera, sunset, diary mallard, coffee). For each interface, participants carried out training trials with the interface, and then testing trials with the Location-only interface (that is, no labels or highlights were visible, although users could still inspect objects by holding their mouse on the object for one second).

Participants saw different datasets for each interface, and order was balanced so that each dataset/interface combination was seen an equal number of times. The dataset was held constant between training and testing.

Procedure

Participants were introduced to the study and asked to carry out a spatial-abilities pretest based on the pair-matching game 'Memory' (also called 'Concentration'). After the pretest, they were shown the different interface conditions. They were shown the Location-only condition last, and were told that they would be using this interface in the testing portion of the study (that is, participants were aware that they would have to eventually perform the task without the navigation aid). Participants were then randomly placed into one of four order groups. They carried out training and testing trials with each interface: five blocks of ten training trials with the interface condition, and then two blocks of ten test trials with the Location-only interface. After each interface condition, participants filled out an effort questionnaire based on the NASA TLX survey [8].

For the prediction interfaces (Predict-50 and Predict-90), participants were told that the computer would try to highlight the next item, but they were not told what the accuracy of the predictor was. They were only told that sometimes the computer would be correct, and sometimes incorrect.

Participants were instructed to work as quickly as possible, but were asked not to complete the testing part of the task simply by clicking on all the objects. If a participant made more than three errors on a trial, the experimenter reminded them to try and be more accurate, and to inspect the objects if necessary. However, participants were also told that if they were fairly confident about an item's location, they were free to click the target without inspecting it first.

Study Design

The study used a within-participants 4x2 factorial design. The factors were:

- Training Interface: Labels, Predict-50, Predict-90, Location-only
- Stage: Training (navigation aids present), Testing (not present)

Participants carried out tasks with all four interfaces; order was balanced such that each aid type was seen in each position an equal number of times.

The sample size (16 people) provides statistical power of 0.8 with a means difference of 300ms, assuming a standard deviation of 400ms (determined from pilot studies) and an alpha level of 0.05. This implies that if there is a true difference of at least 300ms between the conditions, we have an 80% chance of finding a significant difference.

In addition to the effort questionnaires, the system collected time and error data for all trials, both training and testing. With 16 participants, each of whom carried out 50 training and 20 test trials in each of 4 conditions, the system collected data from a total of 4480 trials.

RESULTS

The main questions we wished to answer in our analysis were:

- Are there performance differences between the interfaces during training (i.e., did any the navigation aids assist performance when they were being used);
- Are there performance differences in the testing phase (i.e., does training with a particular interface affect retrieval performance in the Location-only condition, and in particular, does training with certain interfaces hamper testing performance);

We organize the analysis below in terms of our three main measures: completion time, errors, and subjective effort.

Completion Time

A 2x3 ANOVA showed significant main effects for both Interface ($F_{3,45}$=8.55, $p<0.001$) and Stage ($F_{1,15}$=55.79, $p<0.001$). There was also a significant interaction between the factors ($F_{3,45}$=27.38, $p<0.001$). Due to this interaction, we consider training and testing results separately below.

Training (Navigation Aids Present)

For training trials, we found a main effect of Interface ($F_{3,45}$=19.79, $p<0.001$). A follow-up Tukey pairwise test showed that retrieval time in both the Predict-90 and Labels conditions was significantly faster than performance in Locations-only or Predict-50 ($p<0.01$, see Table 1).

	Labels	Predict-50	Predict-90	Locations
Labels		-8.02	-1.92	-6.58
Predict-50	p<0.01		6.10	1.43
Predict-90	n.s.	p<0.01		-4.67
Locations	p<0.01	n.s.	p<0.01	

Table 1. Tukey test results (statistic value above diagonal, p-value below).

Testing (with Location-only interface)

For testing trials, there was no main effect of Interface ($F_{3,45}$=0.90, p=0.45). As shown in Figure 7, there is little difference between the means, although Predict-90 and Labels are slightly faster than Predict-50 and Locations.

Does this analysis mean that there really is no effect of Interface on subsequent performance without the navigation aid? Based on the earlier power analysis, we have an 80% chance of seeing a means difference of 300ms; therefore, while there may be an unseen significant difference of less than 300ms, there is unlikely to be a larger one.

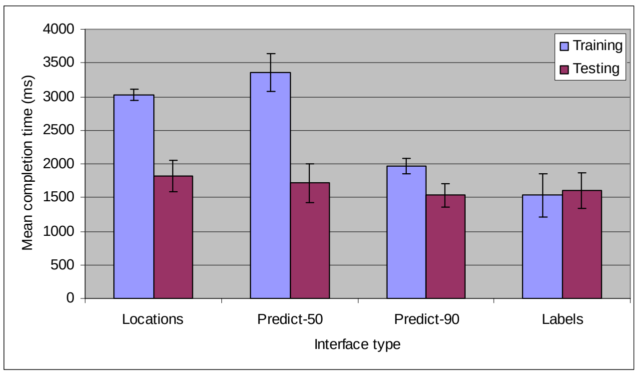

Figure 7. mean completion time per trial with navigation aid present (training) and with aid removed (testing). Error bars show one standard error from the mean.

Retrieval Performance Over Time

The results summarized in Figure 7 show that the best performance was seen with the two interfaces that required the least amount of location memory (Labels and Predict-90). As a post-hoc analysis, we examined the data by trial block (ten trials/block) in order to see how performance changed over time. Figure 8 shows retrieval time for each interface by trial block (five training blocks with the interface, two testing blocks with Locations-only).

There was a significant main effect of block number ($F_{6,90}$=65.62, p<0.001), but there was also a significant interaction between block and Interface ($F_{18,270}$=11.13, p<0.001). As can be seen in Figure 8, different interfaces led to different improvements over time: the two location-dependent interfaces make much greater improvement than either Labels or Predict-90.

We tested each block separately using a Tukey test to determine where there were differences between interfaces. For training trials (all p<0.05):

- the only time Labels and Predict-90 are significantly different is in block 1; Locations and Predict-50 are not significantly different in any block;
- Locations and Predict-50 are significantly slower than the other two interfaces in blocks 1 to 3;
- by block 4, Locations is no longer significantly different from any other interface;
- in block 5, the only difference was that Predict-90 was significantly faster than Predict-50.

For testing trials, there were no differences between interfaces for either block (at p<0.05). However, as shown in Figure 9, retrieval times do increase for the two low-spatial-memory interfaces when the navigation aid is removed, and retrieval times for the high-spatial-memory interfaces do not. In addition, the interface with the least dependence on spatial location memory (Labels) has the larger increase in retrieval time (from 1330ms in block 5 to 2150ms in block6, an increase of 820ms). Retrieval time for these two interfaces has gone down somewhat by the second test block, however.

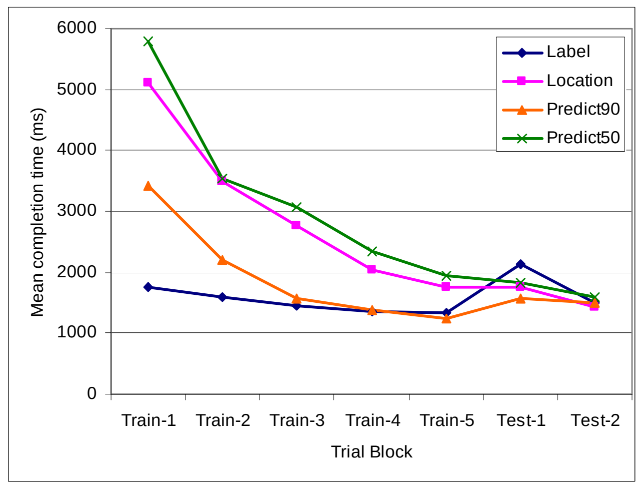

Figure 8. Mean completion times by trial block.

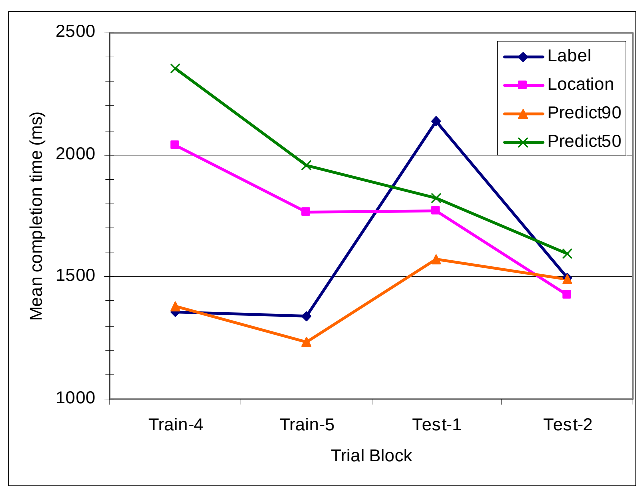

Figure 9. Detail of transition between training (with aid) and testing (without aid).

Errors

Any incorrect selection was counted as an error; the distance from the target was not recorded. A 4x2 ANOVA showed no significant main effect of Interface ($F_{3,45}=1.62$, $p=0.20$), and no main effect of Presence ($F_{1,15}=1.74$, $p=0.21$). There was a clear interaction, however, between the factors ($F_{3,45}=11.29$, $p<0.001$), and we analyse the data separately by Stage below.

Training

For training trials, a main effect of Interface was found ($F_{3,45}=6.26$, $p<0.005$). A subsequent Tukey test showed that all interfaces were significantly different ($p<0.01$) except for the difference between Location and Predict-50. As shown in Figure 10, there are far more errors in the high-spatial-memory conditions (Locations and Predict-50, approximately one error in seven trials; Predict-90, one error in twenty trials; and Labels, one in more than a hundred).

Testing

A main effect of Interface was also found for testing trials ($F_{3,45}=7.64$, $p<0.001$); however, in this case the error rates of the conditions were reversed. Participants made the most errors when they trained with Labels (one error in about eight trials). Participants who trained with Predict-90 and Predict-50 made one error in about fifteen trials, and those with Locations, about one in seventy).

The error rates suggest that accuracy is dependent on the strategy that people use to find the object. When first learning the dataset, people using Locations-only made a large number of errors, and those with Labels made very few. After the labels were removed, the situation was almost perfectly reversed. It should be noted, however, that even the highest of the recorded error rates (about one in seven) might not be a problem in many real-world tasks; in addition, we encouraged participants to select items before the popup label appeared (if they were confident that they knew the location); in a real-world system with a faster popup, some of these errors would be avoided.

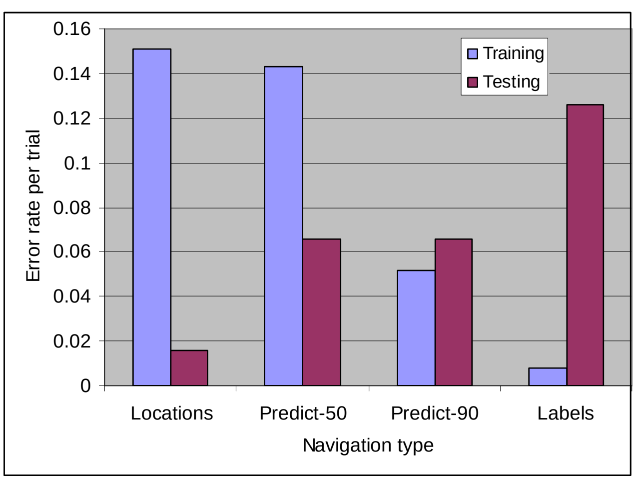

Figure 10. Mean error rates for training and testing.

Error Rate over Time

Figure 11 shows error rates by trial block. As above, we analysed differences within each block. The main results were ($p<0.05$):

- Location and Predict-50 have significantly higher error rates in blocks 1-4 than Labels or Predict-90;
- Predict-90 has significantly more errors than Labels for blocks 1 and 2;
- Labels has a significantly higher error rate than Location for both testing blocks.

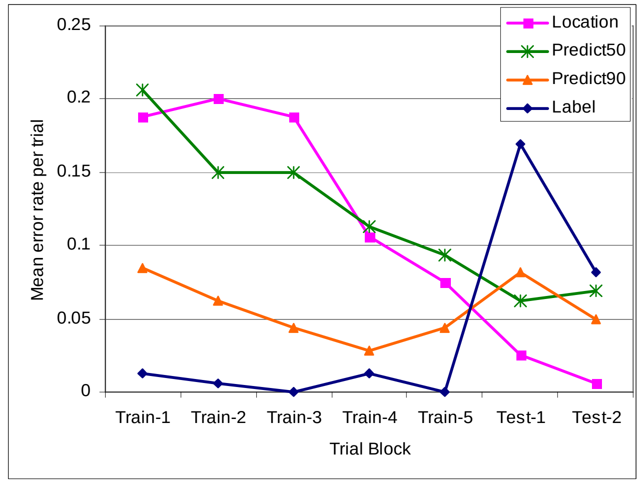

Figure 11. Mean error rates per trial, by block of ten trials.

Subjective Effort

After testing with each interface condition, participants filled out an effort questionnaire based on the NASA Task Load Index [8]. The questionnaire asks participants to rate the task on a scale of one to five in terms of mental demand, physical demand, temporal demand, subjective performance, overall effort, and frustration level.

Results are shown in Figure 12. We analysed each question separately using 4x1 ANOVAs. Two questions showed significant effects of Interface: mental demand (q1) and subjective performance (q4) ($p<0.05$).

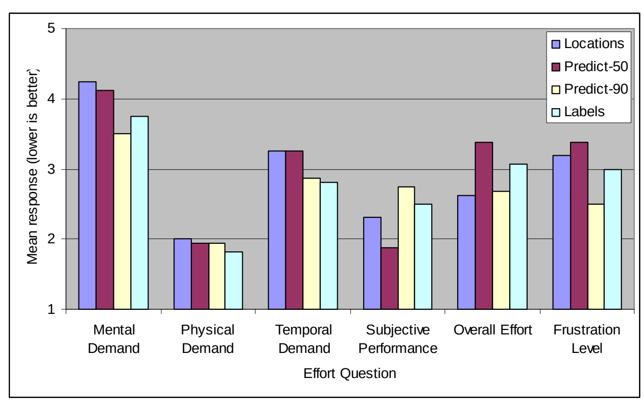

Figure 12. Mean scores for effort questions (lower is better for all questions).

Effects of Prior Spatial Ability and Sex

We tested people for prior location memory using a computational version of the game ‘memory.’ The system presented a 4x5 array of cards, face-down on the screen. Participants turned over cards two at a time, trying to find matching shapes (there were ten pairs in the set). The system recorded the number of card-flips, and we use the total as a rough measure of people’s ability to remember the locations of objects.

We divided participants into two groups based on the median of the sample; we call the group above the median the ‘high-spatial’ group, and those below the median the ‘low-spatial.’ We carried out a post-hoc test using these groups to see whether there were any effects of prior spatial ability on retrieval time in the study.

We found that although the low-spatial group was slightly slower overall (2.36 seconds compared to 2.11 seconds), there was no main effect of spatial ability ($F_{1,14}$=0.59, p=0.45), and no interaction between spatial ability and interface type ($F_{3,42}$=0.19, p=0.91).

We also tested to see if there were differences between the male and female participants, given the prior research into sex differences in spatial abilities (e.g., [18]). No main effect ($F_{1,14}$=0.33, p=0.57) or interaction with interface ($F_{3,42}$=2.45, p=0.07) was found. One data point should be explored further, however – that women were almost a second slower than men in the Predict-50 condition (3.3 seconds vs. 2.4 seconds).

DISCUSSION

The main findings from the study are:

- Some navigation aids (Labels and Predict-90) significantly improved retrieval times during training;
- A less-reliable navigation aid (Predict-50) was not significantly different than no aid at all;
- When the navigation aid is removed, there is little effect on retrieval performance: there is an initial disruption, but performance recovers quickly;
- Users stated that the (effective) navigation aids required the least amount of effort;
- The Labels interface did lead to more errors in testing, particularly immediately after it was turned off.

We draw two main conclusions from this work. First, location learning does appear to come ‘for free’ – people learn locations well, even with training methods that require very little location knowledge. The navigation aids that we tested do not create any kind of ‘speed-dial dependence’ that hampers performance when the aid is removed. Second, there were obvious benefits of navigation aid in the early parts of the task, when the data was not well known. Performance in training trials with Labels and Predict-90 was substantially better (2-4 sec.) than with Location-only.

These results clearly suggest that designers should use navigation aids, without concern that the aid may not always be available. Designers should also consider the efficacy of the navigation aid, since a poor aid may be no different than none at all.

Explanations for results

Here we consider explanations for three results from the study: why Labels and Predict-90 worked well in the training phase; why removing these aids did not affect performance; and why Predict-50 did not improve performance.

First, the two better navigation aids appeared to assist performance in training for exactly the expected reasons: that is, they simplified the user’s task of finding the required item. Labels worked by allowing the user to carry out a visual search on the words, and Predict-90 worked because the predicted item was almost always the right object. Even though some of the highlighted items were wrong, the cost of searching in these cases was more than made up by the advantage of the accurate prediction.

Second, it is clear that people successfully learned the locations, even with the lower-effort interfaces. It appears that simply interacting with the objects provided enough experience for the locations to be remembered. In terms of the different types of location learning discussed earlier, this study involved intentional learning – in that participants knew that they were going to have to perform without the navigation aid. Therefore, we do not know if the location learning was completely incidental; this is a question for further study. We are interested in what part of the interaction results in location learning, and in what would happen with even lower-effort strategies (e.g., Predict-100). We are also interested in whether real-world users think of the possibility that they will not always have access to their navigation aid; that is, whether their location learning is incidental or at least partly intentional in real-world tasks.

Third, our observations of the Predict-50 condition suggests that people did not trust the prediction enough to use it, and depended on their spatial memory instead. Since participants did not know what the accuracy of the predictor was, they may have assumed that it was extremely unreliable, and not worth checking at all. We plan to follow up this result in future work, and examine people’s inspection behaviour more closely with navigation aids of varying accuracy (e.g., where would Predict-70 fit?). It is worth noting, however, that poor prediction did not ever hinder participants; it merely did not help. In addition, in the real world, users will be able to bring contextual information to bear on whether to inspect a highlighted item, reducing the number of times that a user would be ‘fooled’ into checking an incorrect prediction.

Are there advantages to Location-only?

Although there is a substantial cost in the early stages, people did seem to learn the locations better with Locations-only (based on equal retrieval time and lower errors in testing). It is not clear whether people who trained with Labels or Predict-90 would catch up to the Location-only group; more study is needed to determine which training method will lead to optimal performance over the longer term.

Generalization

Several potential issues arise in considering whether our results will generalize to real-world situations:

Other types of navigation aid. Our study tested two types of aid: labels, an explicit representation of the item; and prediction, which showed a highlight of only one item. Other types of aid are similar in many respects to these two (i.e., they are either explicit or predictive), and we expect similar results could be found with user-chosen flags or visit wear.

Larger datasets. Our system used only ten items; real-world datasets involve far more objects. Although the basic principles seen in the study are likely to transfer to larger datasets, further study is required to determine the effects of distracters (e.g., remembering ten items out of a set of dozens or hundreds).

Contextual information. Our study did not provide explicit landmarks, background visual information, or a meaningful task context (where items were grouped by some task-related attribute). Previous research suggests that context and landmarks will improve location learning; it is yet unknown whether these factors will interact with the presence of navigation aids.

Realistic tasks. Two differences between our study task and a real-world task are that a real-world situation would involve more time between retrievals, and that people would bring task knowledge to the retrieval process. Both of these factors should be studied further; however, our experiences suggest that navigation aids could be an even greater help in situations where revisitation is less frequent.

Comparison to previous work

The two sides in the question of navigation aids are represented by two previous studies: Ehret's study of location learning [7], and Czerwinski et al.'s study of the Data Mountain [5]. Here we compare our results to these earlier projects.

Ehret's study. Our results appear to contradict Ehret's hypothesis that location learning is strongly tied to the location knowledge that is required by the retrieval strategy. However, when we look more closely at his results, they do not seem so different from ours. In Ehret's location post-test, only the purely-perceptual colour-matching condition led to reduced location memory; the icon and label conditions were no different from the location-only condition. We did not have a purely perceptual condition similar to Ehret's colour-match. It is possible that there is still a negative effect for extremely low-effort strategies; we plan to study this issue further in future by testing a 'Predict-100' condition.

Czerwinski et al.'s study. Our results agree with those of the followup study of the Data Mountain. Although our experimental setup, task, and amount of training was considerably different, we also found a short-term reduction in retrieval time (but no significant difference) when a navigation aid was removed. In addition, we also found an increase in errors once the navigation aid was removed (in particular, with the Labels condition).

CONCLUSION

Navigation aids are visual markers that assist users in finding desired objects in large datasets. In this paper, we investigated the question of whether, by making items easier to find, navigation aids have a negative effect on location learning. We carried out a study of four different training interfaces that required various amounts of location knowledge for retrieval. The interfaces included one that showed item labels, two that predicted the next item, and one that provided no navigation aid. The better navigation aids made a significant difference during training, but once the navigation aid was removed, there was no significant difference in retrieval performance. We conclude that navigation aids are valuable when the dataset is not well known, and that they do not hinder eventual learning of object locations.

In future work, we plan to test several additional questions raised by this study. First, we will investigate different navigation aids (such as visit wear and bookmarks) and more realistic tasks and datasets. Second, we plan to test the effects of navigation aids over a longer term, and test predictors with different accuracies to find out at what point people stop considering the prediction. Third, we plan to carry out a followup study to find out whether location learning can occur even with perfect prediction.

ACKNOWLEDGMENTS

<removed for review>

REFERENCES

1. Andrade, J., and Meudell, P. (1993). Short Report: Is Spatial Information Encoded Automatically?, *Quarterly Journal of Experimental Psychology*, 46A, 365-375.
2. Baddeley, A. (1990). *Human Memory*. Hove: Lawrence Erlbaum Associates.
3. Baudisch, P., and Rosenholtz, R. (2003). Halo: A Technique for Visualizing Off-Screen Locations, *Proc. ACM CHI 2003*, 481-488.
4. Becker, R., and Cleveland, W. (1987). Brushing Scatterplots, *Technometrics*, 29(2), 127-142.
5. Czerwinski, M., van Dantzich, M., Robertson, G., and Hoffman, H. (1999) The Contribution of Thumbnail Image, Mouse-over Text and Spatial Location Memory to Web Page Retrieval in 3D Viewing, *Proc. IFIP INTERACT 1999*, 163-170.

6. Darken, R., and Sibert, J., Wayfinding in Large-scale Virtual Environments, *Proc. ACM CHI 1996*, 142-150.
7. Ehret, B. (2002). Learning where to look: location learning in graphical user interfaces, *Proc. ACM CHI 2002*, 211-218.
8. Hart, S., and Staveland, L. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research., in Hancock, P., and Meshkati, N., eds., *Human Mental Workload*, Amsterdam: North Holland, 139-183.
9. Hasher, L., and Zacks, R. (1979). Automatic and Effortful Processes in Memory, *Journal of Experimental Psychology: General*, 108, 356-388.
10. Hill, W., Hollan, J. Wroblewski, D. and McCandless, T., Edit Wear and Read Wear. *Proc. ACM CHI 1992*, 3-9.
11. Hite, T. (2003). *Gender, Spatial Learning Trials, and Object Recall*, unpublished M.A. thesis, East Tennessee State University, 2003 (etd-submit.etsu.edu/etd/theses/available/etd-0709103-143240/
12. Jones, W., and Dumais, S., The Spatial Metaphor for User Interfaces: Experimental Tests of References by Location versus Name. *ACM TOIS*, 4, 1986, 42-63.
13. Lewis, J., Rosenholtz, R., Fong, N., and Neumann, U. (2004). VisualIDs: Automatic Distinctive Icons for Desktop Interfaces, *Proc. ACM SIGGraph'04*, 416-423.
14. Lynch, K., *The Image of the City*. Cambridge: MIT Press, 1960.
15. Postma, A., & De Haan, E. (1996). What Was Where? Memory for Object Locations, *Quarterly Journal of Experimental Psychology*, 49A (1), 178-199.
16. Robertson, G., Czerwinski, M., Larson, K., Robbins, D., Thiel, D., and van Dantzich, M., Data Mountain: Using Spatial Memory for Document Management, *Proc. ACM UIST 1998*, 153-162.
17. Siegel, A., and White, S., The Development of Spatial Representation of Large-Scale Environments. In H.W. Reese ed., *Advances in Child Development and Behavior*, New York: Academic Press, 1975.
18. Silverman, I., and Eals, M., Sex differences in spatial abilities: Evolutionary theory and data. In J. Barkow, L. Cosmides, & J. Tooby, eds., *The Adapted Mind*, New York: Oxford Univ. Press, 1992, 533-549.
19. Skopik, A. and Gutwin, C. (2005). Improving Revisitation in Fisheye Views with Visit Wear, *Proc. ACM CHI 2005*, 771-780.
20. Tauscher, L. and Greenberg, S., Revisitation Patterns in World Wide Web Navigation. In *Proc. ACM CHI 1997*, 398-406.
21. Van Asselen, M., Fritschy, E., and Postma, A. (2005), The Influence of Intentional and Incidental Learning on Acquiring Spatial Knowledge During Navigation, *Psychological Research*, preprint available at springerlink.metapress.com/openurl.asp?genre=article&id=doi:10.1007/s00426-004-0199-0
22. Wexelblat, A., & Maes, P. (1999). Footprints: History-Rich Tools for Information Foraging, *Proc. ACM CHI 1999*, 270-277.