Evaluating the Effectiveness of Spatial Memory in 2D and 3D Physical and Virtual Environments

Andy Cockburn and Bruce McKenzie

Human-Computer Interaction Lab
Department of Computer Science
University of Canterbury
Christchurch, New Zealand
+64 3 364 2362 http://www.cosc.canterbury.ac.nz/
{andy, bruce}@cosc.canterbury.ac.nz

ABSTRACT

User interfaces can improve task performance by exploiting the powerful human capabilities for spatial cognition. This opportunity has been demonstrated by many prior experiments. It is tempting to believe that providing greater spatial flexibility—by moving from flat 2D to 3D user interfaces—will further enhance user performance. This paper describes an experiment that investigates the effectiveness of spatial memory in real-world physical models and in equivalent computer-based virtual systems. The different models vary the user's freedom to use depth and perspective in spatial arrangements of images representing web pages. Results show that the subjects' performance deteriorated in both the physical and virtual systems as their freedom to locate items in the third dimension increased. Subjective measures reinforce the performance measures, indicating that users found interfaces with higher dimensions more 'cluttered' and less efficient.

Keywords

Spatial memory, 3D user interfaces, document management, information visualization.

INTRODUCTION

Several experiments have shown that spatial organization of information allows efficient access to items in graphical user interfaces. For example, in evaluating their 3D 'Data Mountain', Czerwinski, van Dantzich, Robertson and Hoffman [4] found that spatial memory allowed rapid access to web-page thumbnails several months after the pages were originally organized. A subsequent study showed no significant difference between subjects' ability to retrieve images in 2D and 3D versions of the Data Mountain [2], but revealed that subjects were surprised at the speed and accuracy of their retrieval.

Spatial memory, then, appears to be a valuable tool in supporting efficient information organization. The questions investigated in this paper concern the effectiveness of spatial memory as interfaces move from 2D toward 3D spatial organizations. The investigation is relevant because there is increasing research and commercial interest in systems that provide 3D interfaces for standard file and document management tasks. Example systems include the Task Gallery [12] and Win3D (www.clockwise3d.com), both of which provide 3D surrogates for the standard 'flat' desktop environment. We are interested in the differences, if any, between the way people interact with equivalent physical and virtual tools that support spatial organizations of information with varying levels of support for a third dimension. The objective of the research is to help improve the design of 3D systems for everyday office tasks.

The interfaces used in the evaluation are all motivated by the Data Mountain [11]. The three physical systems used in the experiment and their virtual equivalents are shown in Figure 1.

The next section describes related work, followed by a description of the six interfaces used in the evaluation. The experimental design is then presented, followed by the results, discussion and conclusions.

RELATED WORK

Two areas of related work are particularly relevant to this investigation. First, there has been extensive prior research comparing the effectiveness of 2D and 3D user interfaces. Second, several researchers have examined the role of spatial memory in predicting user performance with graphical user interfaces and in supporting information retrieval.

Comparing 2D and 3D interfaces

Much of the work comparing 2D and 3D (or 'perspective') visualizations has been conducted within aviation and military domains. Many of the findings are dependent on the precise tasks under analysis. Wickens, Liang, Prevett and Olmos [17], for example, examine navigation on an aircraft landing approach with 2D and 3D displays. Their
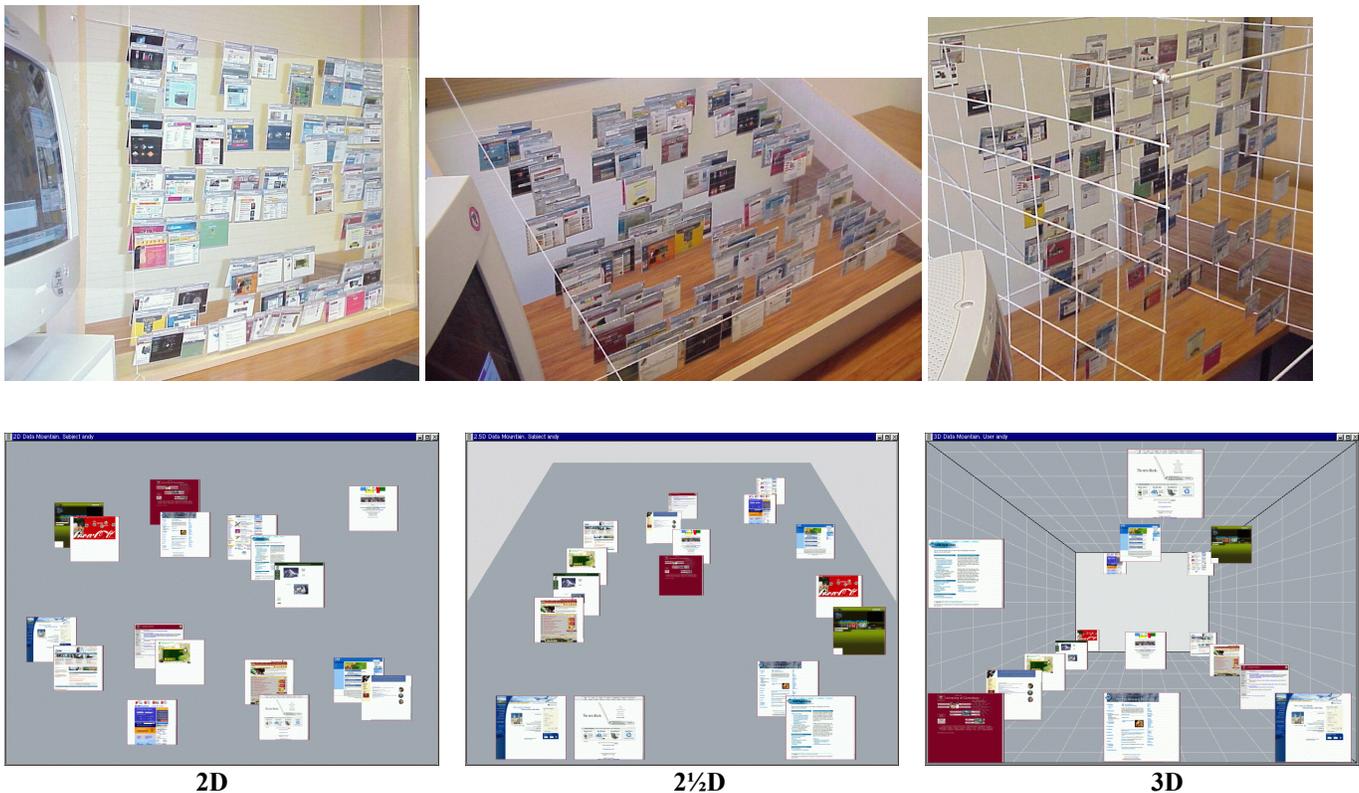
Figure 1: The three physical (top row) and virtual (bottom row) interfaces.

results are mixed, showing that their 3D interface better supported navigation on the lateral axis, but at a substantial cost to performance on the vertical axis. Tests of their subjects' terrain awareness revealed slightly better performance in the 2D condition. The terrain comprehension result contradicts that of St. John, Oonk and Cowen [14], which shows better understanding of terrain shape using a 3D interface. However, another of St. John *et al.*'s tasks showed better understanding of the vertical axis in 2D: subjects were better able to find the highest point on a map in the 2D condition. Further problems with altitude assessment in 3D are also reported in [18]. Delucia [5] shows that 'ground-intercept' information—consisting of a line between objects and the ground—can ease the depth/altitude ambiguity that occurs in 3D visualizations, and Barfield and Rosenberg [1] show that stereoscopic views can also aid altitude assessment.

Outside aviation research there have been several evaluations showing no significant differences between 2D and 3D. Risden, Czerwinski, Munzner and Cook [10] compared 2D and 3D visualizations of web content, finding no significant differences in user performance or satisfaction with the visualizations. Similarly, Cockburn and McKenzie [2] compared 2D and 3D versions of the Data Mountain [11] and found no significant difference in performance, but a significantly higher subjective rating for 3D. Ware and Franck [16] conducted a comprehensive comparison of 2D and various modes of 3D in tasks

involving comprehension of 3D graphs. The 3D conditions reliably outperformed 2D, with the ability to move or rotate the graph proving a valuable feature in 3D. Finally, Smallman, St. John, Oonk and Cowen [13] showed that 2D symbolic representations of military targets allowed faster and more accurate identification than 3D icons.

Wickens, Olmos, Chudy and Davenport [19] provide a fitting summary for prior work on 2D versus 3D evaluations: "whether the benefits of 3D displays outweigh their costs turns out to be a complex issue, depending upon the particular 3D rendering chosen, the nature of the task, and the structure of the information to be displayed." We are not aware of prior comparisons of spatial memory across dimensions and across physical and virtual systems.

Spatial memory and user interfaces

Several papers provide evidence that spatial aptitude is a strong predictor of performance with computer-based user interfaces. The results of Egan and Gomez [6], Gagnon [7], Vicente, Hayes and Williges [15] and Leitheiser and Munro [9] all concur that measures of spatial cognition strongly predict performance with computer interfaces. The interfaces used in their evaluations were a text editor, video games, a hierarchical file structure interface, and file management on the Apple Mac respectively.

The spatial arrangement of web page images provided by the Data Mountain allowed more rapid and accurate retrieval of pages (from sets of 100 pages) than the

'Favorites' mechanism in Microsoft Internet Explorer [3,11]. A subsequent evaluation of the Data Mountain showed that subjects were able to rapidly retrieve pages six months after creating their spatial organization [4]. Furthermore, replacing the thumbnail images with blank outlines did not detrimentally affect retrieval time.

Jones and Dumais [8] provide some cautions on over-reliance on spatial organization. Their evaluation shows that semantic labels provide stronger retrieval cues than spatial organization alone, but indicate that combinations of semantic and spatial organization enhance performance.

2D, 2½D, 3D PHYSICAL AND VIRTUAL INTERFACES

Six interfaces were used in the evaluation—three physical interfaces and three computer-based 'virtual' equivalents. Photos of the three physical interfaces, and screen shots of the three virtual equivalents, are shown in Figure 1. Physical *and* virtual interfaces were used to help ensure that results are indicative of human spatial capabilities rather than some artifact of computer displays. We use the terms 2D, 2½D and 3D to describe the three physical and virtual interfaces. It is important to note that the virtual interfaces provide a fixed perspective view, without enhanced 3D display capabilities such as stereoscopy or head-coupled motion. They provide a platform for investigating fixed perspective spatial memory, and are not intended to test wayfinding or other forms of spatial navigation in virtual worlds.

The 2D interface provides a vertical surface allowing the x and y coordinates for each page to be manipulated. Pages can overlap one another, therefore adding a small element of a third 'depth' dimension. The occlusion algorithm for overlapping pages is deterministically applied so that pages with a lesser value on the y-axis (pages lower in the display) are placed in front of pages with higher y-axis values.

The 2½D interfaces provide a receding inclined plane on which pages can be located. These interfaces are closely analogous to the original '3D' Data Mountain [11]. The x and z coordinates for each page can be manipulated, but the y coordinate cannot (pages hang from the plane in the physical system and stand on the plane in the virtual one). Pages 'recede' in the virtual interface by dynamically reconfiguring the image size, providing a sense of perspective.

The 3D interfaces allow the x, y, and z coordinates for each page to be specified. Details of the physical and virtual implementations of each of these interfaces follow.

The Physical Interfaces

The web page 'thumbnails' in the physical interfaces were 90×90mm photo-quality printed images of pages rendered in Netscape Navigator. The images were mounted on stiff cardboard and covered in clear plastic for protection. The title information for each page was overlaid on top of the Netscape window banner information, allowing a clearly identifiable text title at a distance of approximately two metres. Clips on the back of each card allowed them to hang from the fishing-line used in each of the physical interfaces. The experimenters positioned the cards within the physical interfaces under the subjects' instructions.

The 2D physical interface was constructed from chipboard and horizontal lines of taut fishing-line separated by 2cm in a single vertical plane. String marked the 900×710mm page placement boundaries. The subjects were able to overlap the page cards vertically and horizontally. Overlapping was always achieved so that the top-left corners of the cards, which revealed the page titles, remained visible.

The 2½D physical interface was created by reclining the 2D interface to an angle of 25°. The cards hung vertically off each fishing-line. The string placement boundaries were the same as the 2D interface (900×710mm).

The 3D interface was constructed from painted steel rods and horizontal lines of taut fishing-line placed at 5cm intervals vertically and horizontally. On every second horizontal plane of lines, each second line was removed to create a 'tunnel' of lines to allow the experimenter to place cards within the structure. The 3D structure allowed pages to be placed in a 900×900×750mm x, y and z space. The larger size of the 3D interface on the y-axis (900mm versus 710mm for the 2D interface) was intended to partially compensate for the coarse granularity of placement alternatives on the vertical axis (5 or 10cm versus 2cm in the 2D interface). Cards could be overlapped under similar conditions to those for the 2D and 2½D interfaces.

In all three physical interfaces the fishing-line caused minimal occlusion of pages.

When using the physical interfaces the subjects sat on a height-adjustable chair set approximately 50cm from the front-edge of the interface. This gave an angle at the eye of approximately 84° between the left and right front edge of each interface and approximately 40° at the back of the 3D interface. Head positions were normally approximately mid-height in the 2D interface, towards the top edge of the 2½D interface, and one-third height in the 3D interface.

The subjects used a laser pointer to identify target pages when using the physical interfaces.

The Virtual Interfaces

All three virtual interfaces were written in Tcl/Tk and created windows of 800×600 pixels. The display resolution was 1024x768 with 79 dots-per-inch. Subjects sat approximately 50cm from the screen, giving a horizontal angle at the eye of roughly 30° for the 'front' edge of each interface and approximately 18° and 9° for the back edge of the 2½D and 3D interfaces.

The thumbnails used in all three interfaces were miniaturized rendered web pages. Unlike our earlier experiment [2], the browser's banner information was not included in the thumbnail images because it was reported to
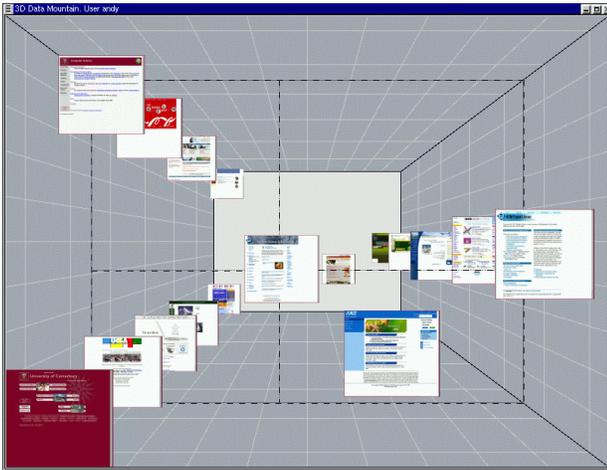
Figure 2: Ground-intercept information for the thumbnail being moved in the 3D virtual interface.

detrimentally affect the subjects' ability to visually identify pages.

In all three interfaces, pressing and holding the right mouse button over any thumbnail magnifies it to 250×250 pixels, and reveals the page title information.

In the 2D interface, all thumbnails are 85×85 pixels, allowing approximately 63 pages to be placed in the display without any overlapping. The occlusion algorithm for overlapping pages is equivalent to that in the physical systems: when pages overlap, those with lesser y-axis values (lower in the display) are placed in front. This improves on the explicit layering model used in the previous experiment [2]. Pages are positioned by dragging with the left mouse button.

The only difference between the 2D and 2½D interface is that it adds perspective to the display. Rather than using a flat plane (with occlusion) for positioning pages, pages are placed on the receding plane. This makes the top corners of the interface unavailable for thumbnail placement. Pages are positioned by dragging with the left mouse button. As pages are 'pushed' up the plane, the thumbnail images diminish from a maximum size of 130×130 pixels at the bottom to a minimum of 40×40 pixels at the top. Seventy-one images can be arranged in the 2½D interface display before overlapping becomes essential. The 2½D occlusion algorithm is identical to that in the 2D interface.

In the 3D interface, the x, y, and z coordinates for each thumbnail can be altered within a virtual 'cube'. Thumbnails are the maximum size of 130×130 pixels at the front of the placement cube, and the minimum size of 40×40 pixels at the back. Sixty-nine images can be arranged in the display before overlapping becomes essential. Dragging with the left mouse button changes the x and y coordinates of thumbnails. Vertically dragging with the middle button changes the z coordinate, effectively 'pushing' pages further away or 'pulling' them closer. To help overcome the problems of depth/altitude ambiguity

Table 1: Experimental design. Subject assignment to conditions.

Density	Physical			Virtual		
	2D	2½D	3D	2D	2½D	3D
Sp (33)	S1-10	S11-20	S21-30	S31-43	S44-56	S57-69
Med (66)	S1-10	S11-20	S21-30	S31-43	S44-56	S57-69
Den (99)	S1-10	S11-20	S21-30	S31-43	S44-56	S57-69

reported in prior work [5,14,17], whenever the user moves a thumbnail, ground-intercept information is revealed: shown in Figure 2. Natural occlusion rules apply, with pages at the front of the cube occluding those further away.

There were no noticeable performance differences between the three virtual interfaces.

EXPERIMENTAL METHOD

The aim of the experiment is to investigate differences in people's ability to use their spatial memory in physical and virtual systems, and to see what effects occur as richer levels of a third dimension are available.

The experimental design, summarized in Table 1, is a 2×3×3 mixed factorial analysis of variance (ANOVA). The between-subjects factor 'realism' has two levels, physical and virtual. The between-subjects factor 'dimension' has three levels: 2D, 2½D and 3D. The final factor 'density' is within-subjects and has three levels: sparse, medium and dense, measured when the interface contains 33, 66 and 99 pages.

Procedure

The evaluation procedure consisted of repeating storage and retrieval exercises for the sparse, medium and dense conditions. Starting with an empty interface, the subjects added 33 pages, one at a time, with freedom to relocate any pages already in the display. The pages were presented in random order in a separate computer-based cueing interface. For the subjects using the physical interfaces, the cueing interface ran on a display on their left (visible in Figure 1). The cueing condition for each page consisted of a magnified thumbnail of the page, the page URL and its title. Subjects were asked to read the title aloud and to ask for clarification if they did not understand the page topic.

In the physical interfaces, the experimenter quickly accessed the correct page-card and showed it to the subject (this took less than two seconds, as the cards were stored in alphabetical order on a desk to the side). The subject then told the experimenter where to place the card, through a combination of pointing, gestures and comments such as "left a bit", "back a bit", "overlapping Greenpeace, there", etc.

To add pages in the virtual interfaces, the subjects pressed an "Add it" button on the cueing interface, and the page was placed at the front-middle-bottom of the display.

Once the thumbnail/card was in the desired location, the subject proceeded to the next image.

The same set of 99 web pages was used for all subjects, although the order of presentation was random for each subject. The main criterion for selecting pages was that the subjects should be roughly familiar with the page topic. Pages included media providers such as `cnn.com`, major international commercial organizations such as `coke.com`, local retail companies, and a variety of university sites.

Having placed and organized the first 33 pages, the subjects responded to two 5-point Likert scale questions: Q1 "It was easy to place the pages" and Q2 "I will be able to quickly find pages" (disagree 1, agree 5).

The retrieval task consisted of finding, as quickly as possible, ten randomly selected pages from the display, one at a time. The cueing condition for each page was identical to that used for storage: the user was shown the magnified image, its URL and title.

In the physical interfaces, the subjects were given a preparatory "three, two, one" countdown prior to displaying the page to be found. A clock on the computer running the cueing interface started as soon as the page to be found was displayed. The experimenter stopped the clock by pressing the space bar as soon as the subject illuminated the target page using a laser pointer. To help the experimenter stop the clock at the correct moment, subjects were encouraged to clearly state when they had located the page with utterances such as "There!" or "Got it". The clock ran a cumulative count, so if the subject identified the wrong page, they were encouraged to continue, and the clock continued running.

In the virtual interfaces software automatically timed each task with the clock stopping when the subject magnified the target page.

A time limit of 100 seconds was used for the search tasks. After the ten searches, subjects responded to two more Likert scale questions: Q3 "I was able to quickly find pages" and Q4 "The display is cluttered".

On completing the retrieval tasks in the sparse condition (33 pages), the subjects proceeded to the medium condition, adding a further 33 pages to the display. They then retrieved ten randomly selected pages from the 66 pages shown. Finally, subjects repeated the tasks for the last 33 pages in the dense condition.

Having completed all tasks, the subjects responded to a final Likert-scale question: Q5 "Overall the interface/structure provides an effective way of organizing and retrieving web pages".

The 69 subjects were volunteer Computer Science students. Thirty were randomly assigned to one of the physical interfaces, giving ten subjects per interface. Thirty-nine were randomly assigned to one of the virtual interfaces, giving 13 subjects per interface. Each evaluation session lasted approximately one hour. The subject's training typically lasted ten minutes and consisted of organizing and retrieving a set of eight training pages (not included in the evaluation set) using the cueing interface. The subjects were advised that the best way to organize pages was to cluster them into groups of related pages. They were given no hints about what groupings would appear. They were also told that they could ask to modify page locations at any time.

RESULTS

Prior to the experiment we had been concerned about the experimental procedure with the physical interfaces. We were worried that we might knock pages off the structure, that subjects might be unable to pinpoint pages with the laser pointer, that they might find issuing directions for page positions difficult, and so on. None of these situations arose, and the subjects quickly became absorbed in their tasks. Several users of the physical 3D interface stated that the system gave them a strong visual effect of pages 'floating in space' in front of them.

Users of the 2D physical interface stayed relatively static in their seat while solving tasks. Users of the 3D interface, particularly those who made substantial use of the z dimension, often used their upper body to move their head up to one foot to the left or right in order to look around occluding pages. Users of the 2½D interface frequently leaned forwards and upward in order to get a better view between pages.

Across the 2070 trials in the six interfaces, the mean time to retrieve pages was low at 4.13 seconds (σ 2.8). There were a total of five trails where the subjects failed to find the page within the 100-second time limit: one in each of the 3D-physical-dense, 2D-virtual-dense, and 3D-virtual-dense conditions, and two in the 3D-virtual-medium condition.

Many subjects commented that they were much faster at retrieving pages than they expected, indicating that their spatial memory was effective, but not trusted.

Retrieval results

The mean times to retrieve pages, across all densities, with the physical and virtual interfaces were 3.5 (σ 1.8) and 4.6 (σ 3.2) seconds. Although this is a significant difference ($F_{1,63}$=6.4, p<.05), it is not particularly meaningful considering the substantially different mechanisms used to interact with the physical and virtual interfaces. These differences include the size of the interfaces (approximately 90×80cm for the physical interfaces versus 11×8cm for the virtual ones), the pointing interface (laser pointer versus cursor), and timing mechanism (human-activated versus automatic).
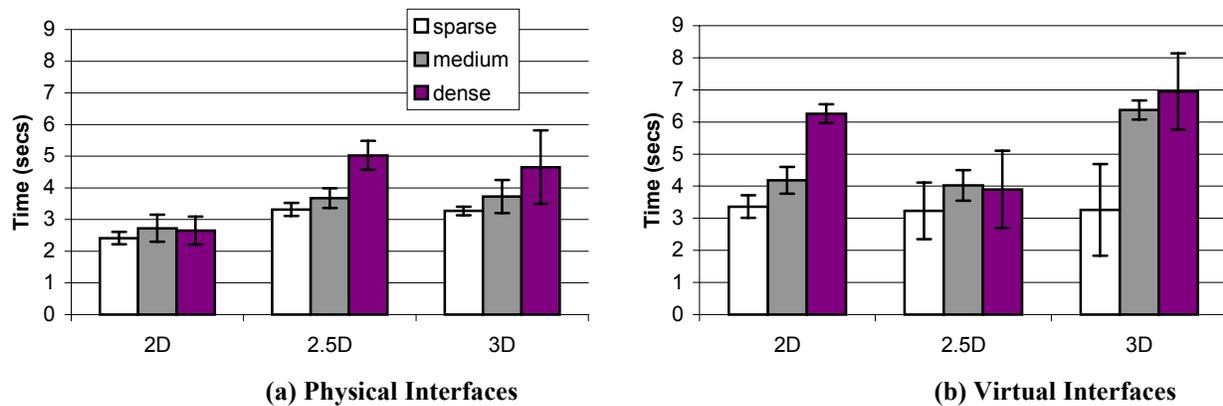
(a) Physical Interfaces   **(b) Virtual Interfaces**
Figure 3: Mean page retrieval times. Error bars show one standard error above and below the mean.

As expected, the means for the three densities were significantly different ($F_{2,126}$=12.8, p<.001) at 3.2 (σ 1.2), 4.2 (σ 2.7) and 5.0 (σ 3.5) seconds for the sparse, medium and dense conditions.

There was a marginal difference between the three levels of dimensionality, with means of 3.7 (σ 2.9), 3.8 (σ 1.5) and 4.8 (σ 3.4) seconds for the 2D, 2½D and 3D interfaces: $F_{2,63}$=2.5, p=.09.

Figure 3a and b summarize the results for the physical and virtual interfaces across dimensions and density. Using the figures to compare results for the 2½D condition with the physical and virtual interfaces reveals the cause of a marginal interaction between factors 'realism' and 'dimensions': $F_{2,63}$=2.6, p=.08. In the physical interface the 2½D interface had the highest mean task time at 4.0 (σ 1.5) seconds, while in the virtual interface it had the lowest mean at 3.7 (σ 1.5) seconds. The relatively poor performance in the physical condition can be explained by the effective reduction in available space for arranging pages. This effect was due to a flaw in the implementation of the 2½D physical interface, described in the discussion section.

Performance with the virtual 2D and 2½D interfaces showed a marked improvement over the previous experiment reported in [2]. Means for the 2D interface decreased from 6.0 (σ 4.8) to 4.6 (σ 3.7) seconds and from 6.8 (σ 5.6) to 3.7 (σ 1.5) seconds with the 2½D interface. These improvements can be attributed to the removal of banner information from the thumbnails and by the modified occlusion behaviour in the 2D interface. Analysis of the results for the virtual 2D and 2½D interfaces confirms the prior result of no difference between the two interfaces: $F_{1,24}$=1.5, p=.2.

A planned comparison of the subjects' retrieval times with the physical 2D and 3D interfaces (Figure 3a) provides an interesting contrast—the 2½D interface is excluded from this comparison because of its implementation limitations described below. Analysis of variance shows a significant difference between the 2D and 3D physical interfaces

$F_{1,18}$=5.5, p<.05, with means of 2.6 (σ 0.6) and 3.6 (σ 1.4) seconds. This is interesting because the 3D interface allowed the subjects to organize images on a 2D plane that was slightly larger than that available in the 2D interface. Two subjects used this strategy to good effect in the 3D interface. Likely reasons for the slower performance with the 3D physical interface are presented in the discussion.

The 3D virtual interface produced the highest mean task completion time at 4.8 (σ 2.9) seconds, although a planned comparison of the three virtual interfaces showed no significant difference between dimensions: $F_{2,36}$=1.3, p=.29. Several subjects complained of 'clutter' and protested that the window was too small when using the 3D virtual interface. The subjective measures reported below reflect this negative reaction. This perception is interesting because the 3D virtual system allowed more non-overlapping pages to be placed in the display (69) than the 2D interface (63), and roughly the same number as the 2½D interface (71).

Subjective measures
The subjects' responses to the Likert scale questions reinforce many of the observed performance comparisons.

Question 1 concerned how easy the subjects found organizing each set of 33 pages in the display. The mean response across all densities with the physical interfaces was 3.6 (σ 1.0) compared to 3.1 (σ 1.1) for the virtual interfaces, yielding a significant difference (Mann-Whitney U-test, U=3794, p<.01). The higher rating for ease of organizing pages in the virtual interfaces is probably due to the larger comparative space available for placing pages without occlusion. This explanation is supported by the subjects' responses to Question 4 "The display is cluttered". Subjects agreed more strongly with this question when using the virtual interfaces (3.7, σ 1.1) than when using physical ones (3.1, σ 1.2): U=3757, p<.01.

In assessing the clutter across dimensions with the physical interfaces, the subjects felt that the 2D interface was significantly less cluttered (2.7, σ 1.2) than the 2½D (3.4, σ 1.0) and 3D (3.4, σ 1.2) interfaces: Kruskal-Wallis H=5.6,

p<.05. There was no significant difference between clutter assessment in the three virtual interfaces (2D 3.9, σ 1.0; 2½D 3.4, σ 1.2; 3D 3.9, σ 1.2): H=0.8, p=.3.

Analysis of responses to Question 2 ("I will be able to quickly find pages") revealed no significant differences between physical and virtual interfaces, or between interfaces with varying numbers of dimensions. Responses to Question 3 ("I was able to quickly find pages"), however, marginally showed worse ratings for the subjects' ability to retrieve pages with the physical interfaces as the number of dimensions increased (2D 4.2, σ 0.8; 2½D 4.0, σ 1.0; 3D 3.7, σ 0.9): H=2.8, p=.06. The virtual interfaces showed no significant differences.

As expected, the subjects became significantly less confident of their ability to quickly find pages (Q2) as density increased, with mean responses of 3.8 (σ 0.7), 3.2 (σ 0.8) and 2.6 (σ 1.1) for the sparse, medium and dense conditions: Friedman χ_r^2 =49.8, p<.001. There was a similar decrease across density in their post-task assessment (Q3), with means of 4.2 (σ .8), 3.9 (σ .9) and 3.4 (σ 1.1): χ_r^2 =16.5, p<.001. In each density, responses to Q2 were significantly lower than their responses to Q3, indicating that the subjects did not trust their spatial memory (Wilcoxon matched-pairs signed-ranks tests). The subjects' comments frequently confirmed this result, with many expressing their surprise that they remembered page locations.

In assessing the overall effectiveness of the interfaces (Q5), responses were significantly higher for the physical interfaces than the virtual ones (physical 4.0, σ 0.9; virtual 3.3, σ 0.9): Mann-Whitney U=341, p<.01. Furthermore, responses with the physical interfaces reliably decreased as the number of dimensions increased (2D 4.5, σ 0.7; 2½D 3.9, σ 0.9; 3D 3.7, σ 0.8): Kruskal-Wallis H=4.4, p=.02. Responses across the three virtual interfaces were noticeably lower (worse) than the physical interfaces, but were not reliably different from each other (2D 3.4, σ 0.7; 2½D 3.8, σ 1.0; 3D 2.8, σ 0.9): H=.03, p=.8.

DISCUSSION

To summarize the results, the time taken to retrieve pages significantly increased through the 2D, 2½D and 3D interfaces. Similarly, the subjects' assessment of the effectiveness of the interfaces decreased through the 2D, 2½D and 3D conditions. As expected, performance deteriorated as the number of pages in the displays increased.

The relatively poor performance with the 3D physical interface is particularly interesting. The 3D interface essentially offers a large superset of the page organization possibilities with the 2D interface, yet the additional flexibility provided by the third dimension appeared to result in confusion for the subjects. Many times we observed subjects in the 3D physical condition uttering statements to the effect of "where did I put that group".

Equivalent statements were much less common in the other interface conditions. Prior to the evaluation we had suspected that problems caused by occlusion might inhibit efficient use of the 3D physical interface. During the evaluation, however, this problem occurred rarely compared to the frequency of forgotten page locations.

We suspect that the results reported for the physical 2½D interface are artificially poor. Its implementation simply involved reclining the 2D interface to an angle of 25°, maintaining the original 2D organization space of 900×710mm. The angle at the eye between the bottom and top of the 2½D organization is therefore substantially less than that for the 2D interface. In the virtual 2D and 2½D interfaces, however, the angle at the eye is the same. One observed effect of the reduced angle of display space in the 2½D physical interface was the frequency with which users leaned forward to gain a perspective above the interface, increasing the angular size of the display, and improving the view between cards. For these reasons, the relatively poor performance with the 2½D virtual interface must be treated with suspicion.

There is one further concern about the physical interfaces. Although the fishing-line was successful at minimizing occlusion, it had two effects on the way the physical systems were used. First, fishing-line, like clothesline, provides a natural affordance of the way it should be used: items hang along it. Most subjects using the physical interfaces relied heavily on a horizontal grouping arrangement for grouping related pages along the lines. With the virtual interfaces, in contrast, grouping arrangements were more commonly achieved through co-location clusters, vertical, and back-to-front (in depth) arrangements.

The second problem caused by the use of fishing-line is that it creates discrete placement locations on the y and z axes. In the 2D and 2½D environments, cards cannot be placed less than 2cm vertically apart, and in the 3D environment they can be no closer than 5cm on the y and z axes. We doubt that this had a major impact on the results because almost all of the subjects using vertical arrangements in the 2D environment requested a minimal vertical separation of 'two strings' (4cm) rather than the 2cm minimum that the interface could support.

CONCLUSIONS

Human capabilities for spatial cognition can allow rapid information retrieval in graphical user interfaces. Several prior experiments have shown that spatial organizations of information enable users to access data items surprisingly quickly. Indicative of these results is the statement by Czerwinski, Dantzich, Robertson, and Hoffman [4] that "3D visualization techniques … can lead to improved user memory for where favorite or frequently used information is stored".

This paper described an evaluation that investigates the contribution of the third dimension in supporting effective use of spatial memory. Three physical models were constructed, with each providing a different constraint on the user's ability to locate cards representing web pages in the depth/perspective dimension. The interfaces were named 2D, 2½D and 3D. Equivalent fixed-perspective virtual systems were also constructed. Both physical and virtual systems were used to help ensure that the results are indicative of human spatial memory, rather than an artifact of human interaction with computer displays.

Results show that our subjects' ability to quickly locate web page images deteriorated as their freedom to use the third dimension increased. Their subjective responses also indicated that they found the 3D interfaces more cluttered and less efficient.

Spatial memory clearly provides an effective aid to information retrieval, but we are skeptical of the role that 3D plays in aiding rapid retrieval of data items from static-perspective spatial organizations. Our results indicate that for relatively sparse information retrieval tasks (up to 99 data items), 3D hinders retrieval. Future work will investigate whether 3D spatial arrangements allow more effective retrieval than a series of 2D planes for larger data sets.

REFERENCES

1. Barfield, W. and Rosenberg, C. Judgments of Azimuth and Elevation as a Function of Monoscopic and Binocular Depth Cues Using a Perspective Display. *Human Factors 37, 1,* 173-181. 1995.

2. Cockburn, A. and McKenzie, B. 3D or Not 3D? Evaluating the Effect of the Third Dimension in a Document Management System. In *Proceedings of CHI'2001*, Seattle, April 2001, ACM Press, 434-441.

3. Czerwinski, M., Dumais, S., Robertson, G., Dziadosz, S., Tiernan, S., and van Dantzich, M. Visualizing Implicit Queries for Information Management and Retrieval. In *Proceedings of CHI'99*, 1999, 560-567.

4. Czerwinski, M., van Dantzich, M, Robertson, G., and Hoffman, H. The Contribution of Thumbnail Image, Mouse-over Test and Spatial Location Memory to Web Page Retrieval in 3D. *Proc. INTERACT'*1999. 163-170.

5. Delucia, P. Effects of Pictorial Relative Size and Group-Intercept Information on Judgments about Potential Collisions in Perspective Displays. *Human Factors 37, 3,* 528-538. 1995.

6. Egan, D. and Gomez, M. Assaying, Isolating, and Accommodating Individual Differences in Learning a Complex Skill. In *Dillon, R (ed): Individual Differences in Cognition.* NY: Academic Press, 173-217. 1985.

7. Gagnon, D. Videogames and Spatial Skills: An Exploratory Study. *Educational Communication and Technology 33,* 4, 263-275, 1985.

8. Jones, WP. and Dumais, ST. The Spatial Metaphor for User Interfaces: Experimental Tests of Reference by Location versus Name. *ACM Transactions on Office Information Systems 4, 1,* 42-63. 1986.

9. Leitheiser, B. and Munro, D. An Experimental Study of the Relationship Between Spatial Ability and the Learning of a Graphical User Interface. In *Proceedings of the Inaugural Americas Conference on Information Systems,* 1995. hsb.Baylor.edu/ramsower/acis/papers/leitheis.htm

10. Risden, K., Czerwinski, M., Munzner, T. and Cook, D. An Initial Examination of Ease of Use for 2D and 3D Information Visualizations of Web Content. *Int. J. Human-Computer Studies, 53,* 695-714. 2000.

11. Robertson, G., Czerwinski, M., Larson, K., Robbins, D., Thiel, D., and van Dantzich, M. Data Mountain: Using Spatial Memory for Document Management. In *Proceedings of UIST'98*, 1998, ACM Press. 153-162.

12. Robertson, G., van Dantzich, M., Czerwinski, M., Hinkley, K., Thiel, D., Robbins, D., Risden, K. and Gorokhovsky, V. The Task Gallery: A 3D Window Manager. In *Proceedings of CHI'2000*, The Hague, April 2000, ACM Press. 494-501.

13. Smallman, H., St. John, M, Oonk, H. and Cowen, M. Track Recognition Using Two-Dimensional Symbols or Three-Dimensional Realistic Icons. *SPAWAR Systems Center Technical Report Number 1815.* 2000. www.nosc.mil/sti/publications/pubs/tr/1818.

14. St. John, M., Oonk, H. and Cowen, M. Using Two-Dimensional and Perspective Views of Terrain. *SPAWAR Systems Center Technical Report Number 1815.* 2000. www.nosc.mil/sti/publications/pubs/tr/1815.

15. Vicente, K., Hayes, B. and Williges, R. Assaying and Isolating Individual Differences in Searching a Hierarchical File System. *Human Factors 29,* 3, 349-359, 1987.

16. Ware, C., and Franck, G. Evaluating Stereo and Motion Cues for Visualizing Information Nets in Three Dimensions. *ACM Trans. Graph. 15,* 2, 121-139, 1996.

17. Wickens, C., Liang, C., Prevett, T. and Olmos, O. Egocentric and Exocentric Displays for Terminal Area Navigation. In *Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting,* 1995, 16-20.

18. Wickens, C., Miller, S. and Tham, M. The Implications of Data-Link for Representing Pilot Request Information on 2D and 3D Air Traffic Control Displays. In *Proc. Human Factors and Ergonomics Society 38th Annual Meeting,* 1995, 61-65.

19. Wickens, C., Olmos, O., Chudy, A. and Davenport, C. Aviation Display Support For Situation Awareness. *Research Report ARL-97-10/LOGICON-97-2.* 1997. www.aviation.uiuc.edu/new/html/ARL/TechPdf/97-10.pdf