Physically Large Displays Improve Performance on Spatial Tasks

DESNEY S. TAN
Microsoft Research
DARREN GERGLE
Northwestern University
and
PETER SCUPELLI and RANDY PAUSCH
Carnegie Mellon University

Large wall-sized displays are becoming prevalent. Although researchers have articulated qualitative benefits of group work on large displays, little work has been done to quantify the benefits for individual users. In this article we present four experiments comparing the performance of users working on a large projected wall display to that of users working on a standard desktop monitor. In these experiments, we held the visual angle constant by adjusting the viewing distance to each of the displays. Results from the first two experiments suggest that physically large displays, even when viewed at identical visual angles as smaller ones, help users perform better on mental rotation tasks. We show through the experiments how these results may be attributed, at least in part, to large displays immersing users within the problem space and biasing them into using more efficient cognitive strategies. In the latter two experiments, we extend these results, showing the presence of these effects with more complex tasks, such as 3D navigation and mental map formation and memory. Results further show that the effects of physical display size are independent of other factors that may induce immersion, such as interactivity and mental aids within the virtual environments. We conclude with a general discussion of the findings and possibilities for future work.

Categories and Subject Descriptors: H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Screen design, user-centered design, graphical user interfaces*; J.4 [**Social and Behavioral Sciences**]—*Psychology*

General Terms: Design, Experimentation, Human Factors, Performance

Additional Key Words and Phrases: Large display, field of view, visual angle, spatial task, immersion, presence, 3D navigation, mental map formation, human memory

1. INTRODUCTION

Even though we have experience in designing both real and virtual worlds, Ishii and Ullmer [1997] observe that the two worlds remain largely disjoint and that there exists "a great divide between the worlds of bits and atoms." In their work, they identify input devices as bridges that serve to connect the two worlds. They focus on understanding how physical objects and architectural surfaces can be used to control digital objects in the virtual world. Using their tangible interfaces, they attempt to build computing environments that support human thought and action.

However, little effort has been spent on understanding the design of the physical computer and its associated display devices [Buxton 2001]. Most work in this area has focused on pragmatic issues surrounding the changing form factors of displays, but few researchers have devoted much attention to understanding how physical affordances of these displays fundamentally affect human perception and thought. As such, design principles have been uniformly applied across a variety of display devices that offer different cognitive and social affordances.

With recent advances in technology, large wall-sized displays are becoming prevalent. Although many researchers have articulated qualitative benefits of group work on large displays (e.g. Swaminathan and Sato [1997]), much less has been done to systematically quantify and exploit these benefits for individual users. Furthermore, within the work aimed at quantifying benefits of large displays, little has been done to understand physical size as an important display characteristic that affects task performance.

In this article, we describe a series of experiments comparing the performance of users working on a large projected wall display to that of users working on a standard desktop monitor. Because we were interested in isolating the effects of physical size, we kept the visual angle subtended from the user to each of the two displays constant by adjusting the viewing distances appropriately (see Figure 1). We also held other factors such as resolution, refresh rate, color, brightness, contrast, and content as constant as possible across displays. Since the information content shown by each of the displays was equivalent, it would be reasonable to expect that there would be no difference in performance on one display or the other. However, we will show that this is not the case, and that physical size is indeed an important display characteristic that must be considered as we craft our display systems.

Results suggest that physically large displays, even at identical visual angles as small displays, increase performance on spatial tasks such as 3D navigation as well as mental map formation and memory. We show through the experiments, how these results might be attributed, at least in part, to large displays immersing users and biasing them into adopting more efficient cognitive strategies. Furthermore, the effects caused by physically large displays seem to be independent of other factors that may induce immersion or increase performance. For example, even though interactivity and mental aids such as distinct landmarks and rich textures within virtual worlds increase task performance on the tasks tested, they did not affect the benefits that large displays offer to users.
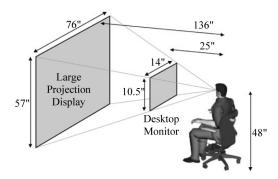
Fig. 1.   Basic experimental setup maintaining visual angles between the small and the large displays by adjusting the distance appropriately.

2. RELATED WORK

In constructing complex workspaces, researchers have pursued the use of large displays for collaborative tasks [Chou et al. 2001; Elrod et al. 1992; Raskar et al. 1998; Streitz et al. 1999; Tani et al. 1994]. Large displays in these settings are easy for all users to see and interact with [Guimbretière 2002], providing a conduit for social interaction. Some of these researchers have begun to document performance increases for groups working on large displays [Dudfield et al. 2001].

While much work has focused on collaboration, less has been done to design for and objectively measure individual gains on large displays. To this end, researchers have explored the use of large displays as a means to provide contextual information to the individual. For example, Baudisch et al. [2002] provide a large low-resolution overview of the working context around a smaller high-resolution focal screen.

Other researchers have realized that large displays may afford users a greater sense of presence, which may benefit performance of certain tasks. Slater and Usoh [1993] define presence as "a state of consciousness, the (psychological) sense of being in the virtual environment." They distinguish it from immersion, which they define to be an objective description of the technology, describing "the extent to which computer displays are capable of delivering . . . illusion of reality to the senses of the human participant." In most current models, the sense of presence is seen as the direct outcome of immersion.

The more inclusive, extensive, surrounding, and vivid the display, the higher the potential of presence [Bystrom et al. 1999]. In fact, when users are present in Virtual Environments (VEs), the location of their physical bodies is often construed as being contained within that space rather than looking at it from the outside. It is in this state that users are most effective in VEs. Tan et al. [2001] utilize large peripheral projection displays to show different scenes of distinct 'places' that the user can use as cues to remember more information. They claim that the greater the sense of presence invoked in the user by the large display, the better the memory for learned information. They do not,

however, articulate explanations for the increased sense of presence on the large display. We discuss several factors that may cause this effect.

One of these factors is field of view (FOV). Large displays are not often placed at a distance that is proportional to their increase in size over small displays. Due to space constraints, they are typically relatively closer and cast a larger retinal image, thus offering a wider FOV. It is generally agreed that wider FOVs can increase "immersion" in VEs [Lin et al. 2002; Prothero and Hoffman 1995]. Researchers in the entertainment industry have reported that larger displays filling a wider FOV can increase the level of involvement experienced by users [Childs 1988]. Czerwinski et al. [2002] report evidence that a wider field of view offered by a large display leads to an increased sense of presence and improved performance in 3D navigation tasks, especially for females. They document prior literature suggesting that restricting FOV leads to negative impacts on perceptual, visual, and motor performance in various tasks, possibly because users find it difficult to transfer real world experience and cognition into the VE. Arthur [2000] provides a comprehensive review of the effects of FOV on task performance, especially as carried out in head-mounted displays.

Despite the large amount of work done in comparing FOVs, few researchers have isolated the effects of physical size and distance on task performance or the sense of presence. To examine the psychophysical effects of distance and size, Chapanis and Scarpa [1967] conducted experiments comparing the readability of physical dials at different distances. They used dials of different sizes and markings that were proportional to the viewing distance so as to keep visual angles constant. Surprisingly, they found that beyond 28 inches away, dials adjusted to subtend the same visual angle were read more easily at greater distances. The effects they found were, however, relatively small.

In a more recent study, Patrick et al. [2000] compared various display technologies, with comparable visual angles, and their effects on the spatial information users acquired by navigating through a VE. They found that while users performed significantly worse in forming cognitive maps and remembering the environment on a desktop monitor, they performed no differently using a head-mounted display or a large projection display. They attributed part of this effect to a higher level of presence afforded by the size of the projection display, which compensated for the immersion afforded by the head tracking. In our work, we further explore the effects of display size and distance, with constant FOV or visual angle, on users' sense of presence and performance on various tasks.

In our work, we chose to explore mental rotation tasks because presentation, degree of immersion, and level of performance have been extensively measured for these tasks. In their work, Suzuki and Nakata [1988] had students perform a mental rotation task similar to that of Shepard and Metzler [1971]. Users were asked to judge whether pairs of figures, each of which had been rotated to different degrees, were identical in shape or not. They found, as Shepard and Metzler did, that mean reaction times increased linearly with the angular difference between figures. They also discovered that wider visual angles, corresponding to wider retinal size, of the objects slowed the speed of rotation. However, in this study, viewing distance, given constant visual angle, did not seem to affect reaction times.
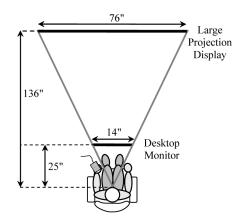
Fig. 2. Top view schematic of the experimental setup. We maintained constant visual angles by varying display size and distance accordingly.

Building on this work, Wraga et al. [2000] measured spatial knowledge by the time it took users to update their orientation after changing it. Results showed that users were faster at spatial updating when they imagined rotating themselves in the environment rather than when rotating the environment around themselves. Carpenter and Proffitt [2001] extended these findings by examining egocentric rotations in each of the three possible rotation planes. They replicated the finding that egocentric rotation, or rotating one's self, was faster than exocentric rotation, rotating the environment, but only for planes in which users had experience rotating or locomoting. Tlauka [2002] found similar results by comparing rotations of images presented horizontally or vertically.

Despite the deep understanding this body of literature offers, there seems to be a gap in work isolating the effects of display size and distance, given a constant visual angle, for performance on tasks. Because of the emergence of large displays in the workplace and in consideration of everyday desktop computing tasks, we decided to evaluate how display size affects performance on spatial orientation and reading comprehension tasks.

3. GENERAL EXPERIMENTAL SETUP

3.1 Equipment

We used two displays for each of the experiments, an Eiki Powerhouse One LCD projector and a standard-sized desktop monitor. In the first two experiments, we used an 18″ Sony Trinitron E400 CRT monitor as the desktop monitor. In the other experiments, we replaced this with an 18″ NEC MultiSync 1810X LCD monitor. All displays ran at a resolution of 1024×768, updated at a rate of 60 Hz, and were calibrated to be of roughly equivalent brightness and contrast. We mounted the projector from the ceiling and projected onto a white wall. The image projected on the wall was 76″ wide by 57″ tall (see Figure 2).

The image on the monitor was 14″ wide by 10.5″ tall. We set the two displays up so that when either display was viewed from a specific spot in the room, the visual angle and hence the size of the retinal image, would be identical. We assumed a comfortable viewing distance of 25″ for the monitor. In order to get an image of identical perceived size, the projection was set up to be 136″ away from the user. The center points of all displays were set to be at seated eye-height, approximately 48″ above the ground.

Since the environmental context around each display could potentially affect users, we decided to keep the context as constant as possible by only moving the displays within the environment rather than having the user turn to face a different display with different environmental context. Hence, we carefully marked the position of the monitor so that it could be moved in and out as necessary.

We ran the exploratory and first two experiments on a single 800 MHz Dell computer equipped with a dual headed nVidia GeForce2 MX graphics card. We controlled the activation and deactivation of the displays using the Windows 2000 multiple monitor API so that only one display was active at any given time. For these experiments, the user provided input using an IBM USB numeric keypad with keys we had marked for the experiment. We ran the latter two experiments on a 1.33 GHz computer with a GeForce4 MX graphics card. The virtual environments updated at 60 frames per second. We used a switchbox to send the graphics output to only one of the displays at any given time. The user provided input with the control stick and trigger button on a Radioshack 26-444 joystick.

3.2 Keeping Color, Brightness, Contrast Constant

We did several things to equate display characteristics such as color, brightness, and contrast across the various displays. Initially, we used a spectral radiometer and a colorimeter to measure the spectral distribution of the light coming off the displays as well as the tristimulus values of this distribution when various images were displayed. Unfortunately, as observed by MacIntyre and Cowan [1992], calibration done to an exact radiometric or colorimetric standard is both expensive and laborious. This is especially true of our setup, in which we were trying to calibrate different display technologies. Calibration is further complicated by human visual phenomena such as light, dark, chromatic, or transient adaptations [Milner and Goodale 1996].

To confirm our calibrations, we took Tjan's [1996] view that a "human observer is always needed to carry out a color matching experiment." In fact, we assumed this to be the case for brightness and contrast as well. After calibrating the displays, we had groups of people view the two sets of displays. With questions such as "which screen do you think is brighter?" or "which screen has better contrast?" we were able to confirm that display settings were as close as we could get them. We iterated this process until users could not make these distinctions between the displays. It is also worth noting that the quality of the large projection display was probably poorer than that of the desktop monitor in all these regards. There is little reason to believe

that the degraded quality would elicit any of the effects that we saw in the experiments.

3.3 Keeping Users' Heads Still

Another concern with the setup was that the visual angle calculations were only valid for a single point in the room. This meant that if users moved their heads from that point, the visual angles were no longer maintained between the two displays. This would cause complication in interpreting results. Even though the most controlled solution would have been to fasten the user's head in place to prevent any movement, we decided against this because it would make the experiment both uncomfortable and unrealistic.

Instead, we marked the spot around which the user's eyes should have been centered by stretching fishing line from two stands, one on either side of the user. A mark in the center of the line indicated the exact spot in the room where the retinal images would be of identical size. For each user, we adjusted the chair so that they were seated comfortably with their eyes as close to the spot as possible and told them not to further move the chair. We then removed the fishing line. In the rare case where users moved their heads or chair too much during the study, we readjusted their position before proceeding. In the initial pilot testing we found that the range of motion was rarely more than 2″ to 3″ in any direction.

At various stages in this work, we also ran informal tests to validate experimental results when users' eyes were either a little too close or too far from the desired point in the room and saw similar effects to those observed in the experiments. Hence, we are fairly confident that the small head movements permitted within the setup did not directly account for the effects seen across the experiments, and we were able to run the studies in a more realistic task environment that allows us to make claims with a greater degree of external validity.

4. EXPLORATORY EXPERIMENTS

In an early set of experiments we showed that although there were no observable differences on a reading task, users performed about 26% better on a spatial orientation task done on the large display as compared to a smaller one, even when we held visual angles and other display characteristics constant [Tan et al. 2003].

We used the Guilford-Zimmerman Spatial Orientation test [Guilford and Zimmerman 1948]. Results from this test have been shown to correlate highly with wayfinding ability [Infield 1991]. Each question in this test contained two pictures seen from the prow, or front, of a boat along with a multiple choice answer key (see Figure 3). The user was asked to imagine that each picture was taken with a camera fastened rigidly to the boat so that the camera bobbed up and down, slanted, and turned with the boat. First, the user looked at the top picture to see where the boat was initially heading. This heading is represented by the dot in the answer key. Next, the user looked at the bottom picture and determined the change in orientation of the boat. The line in each of the possible
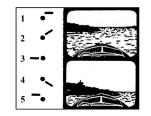
Fig. 3.   Sample question from the Guilford-Zimmerman Spatial Orientation test. The correct answer for this question is option 5.

answers represents the new orientation of the boat relative to the previous heading. Finally, the user selected the answer with the number keys, confirmed the answer with the enter key, and proceeded to the next question. Users had 5 minutes to answer 30 questions in each section, and were told to work as quickly and accurately as possible.

The fact that we found differences in the spatial task but not the reading comprehension task led us to believe that there may be an interaction between the task and the display size. We hypothesized that the performance difference on the spatial orientation task was due to the way the image was perceived and thus the strategy with which users performed the task. In pilot studies, we tried using questionnaires as well as structured interviews to determine the strategy users employed, but found these methods to be inconclusive. Users were not able, either implicitly or explicitly, to articulate their cognitive strategy. Hence, we designed Experiment 1 to explore this more deeply.

5. EXPERIMENT 1: LARGE DISPLAYS BIAS USERS INTO EGOCENTRIC STRATEGIES

One explanation that accounts for performance differences in spatial orientation tasks is the choice of cognitive coordinate systems used to perform the task. This choice usually has implications on the particular strategy and hence the efficiency of performing the task. Just and Carpenter [1985] propose two strategies that might be used to perform the Guilford-Zimmerman test: an egocentric strategy and an exocentric one. Users performing the task egocentrically take a first-person view and imagine rotating their bodies within the environment. Users performing the task exocentrically take a third-person view and imagine objects rotating around each other in space. There is reasonable evidence in psychology research suggesting that egocentric strategies are more efficient for real world tasks (e.g. Carpenter and Proffitt [2001]). Hence,

> *Hypothesis 1a: Simple instructions and training prior to the test are sufficient to bias users into adopting either the egocentric strategy or the exocentric one when they perform the task.*

> *Hypothesis 1b: The egocentric strategy is more efficient than the exocentric one for this spatial orientation task.*

The instructions for the Guilford-Zimmerman test are carefully worded so as not to bias strategy choice one way or another. This allows users to either

Fig. 4. (left) Numeric keypad input device used in the first three experiments. User working on the small (center) and large (right) displays.

imagine themselves on the boat looking through the camera as the boat moves within the environment (egocentric), or outside the environment as the boat rotates within it (exocentric). We believed that as users became more immersed in the task on the large display they were more likely to adopt the egocentric strategy. Since egocentric rotations have been shown to be quicker, this could explain the performance increase we observed on the large display. Thus,

> *Hypothesis 1c: With no explicit strategy provided, display size automatically biases users into adopting one or the other of the strategies. Small displays bias users into adopting an exocentric strategy, and large displays bias users into adopting an egocentric strategy.*

5.1 Participants

Forty-two (18 female) college students from the Pittsburgh area, who did not participate in the exploratory experiments, participated in this one. We screened users to have normal or corrected-to-normal eyesight. The average age of users was 21.8 (21.7 for males, 22.2 for females), ranging from 18 to 35 years of age. The experiment took about an hour and users were paid for their participation.

5.2 Procedure

After users filled out a background survey, we gave them the numeric keypad and had them sit comfortably in the chair (see Figure 4). As previously described, we adjusted the height and position of their chair so that the center of their eyes was as close to the marked fishing line as possible. Once they were viewing the displays from the spot in the room that provided retinal images of identical size, we removed the fishing line. At this point, we instructed users not to further adjust the chair or move it around.

Instructions for the original Guilford-Zimmerman test did not explicitly bias a user into any particular strategy. It describes a boat within a scene, with no indication of the user's place within this environment. The instructions tell users only to "note how the position of the boat has changed in the second picture in relation to the original position in the first picture." From this instruction set, we created two others, one that intentionally biased users into an egocentric strategy and another that biased them into an exocentric

strategy. The egocentric instructions describe a scene in which users are asked to imagine themselves physically on the boat as it moves within the environment. The egocentric instructions now read, "You are standing on top of a boat that is on the movie set. The crew is moving the boat as you are on the boat. Two pictures are taken, one before the boat moves and one after. In each item you are to note how the position of the tip of the boat has changed in relation to the painted backdrop." The exocentric instructions describe the boat as a rigid prop mounted to the ground with the scene on a backdrop that is moving with respect to the boat. Hence, the instructions given are, "You are standing on top of a boat firmly attached to the floor of the movie set. The crew is moving a painted backdrop on the set. Two pictures are taken, one before the painted backdrop is moved and one after. In each item you are to note how the position of the painted set backdrop has changed relative to the tip of the boat on the movie set." For the entire instruction and stimulus set, see Tan [2005]. After balancing for Gender, we randomly assigned each participant to one of the three Instruction Types: Egocentric Instructions, Exocentric Instructions, or Original Guilford-Zimmerman Instructions.
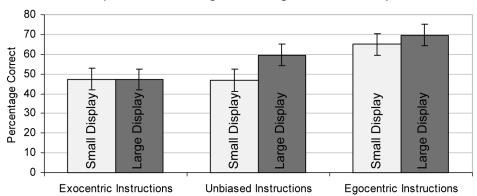
We gave users paper-based instructions appropriate for the condition they were in. They then tried three practice questions. For these questions, the system provided users with immediate feedback explaining the correct answers. After they had completed the practice questions, users performed the test on both the small and the large display, which we will refer to as the Display Size manipulation. The order of Display Size was counterbalanced across users. Users were not given feedback for the test questions. The 60 test questions were randomized and broken into two sets. Users had 5 minutes to answer 30 questions in each of the two conditions, and were told to work as quickly and accurately as possible. Users had a 30 second rest interval between each condition.

After users completed the tests, they filled out a questionnaire indicating their preference for the conditions in each of the tasks. They were also encouraged to comment on their opinion of the displays.

5.3 Results

First we present performance results for the spatial orientation task, and then we examine the preference data.

5.3.1 *Effects of Strategies on Task Performance.* We analyzed data for the spatial orientation task at the summary level. The dependent variable was the percentage of correct responses (number correct/number attempted). Time differences between different Display Sizes were not significantly different and were therefore dropped from the final models. Levels of significance did not change either way. We analyzed the percentage of correct responses with a 2 (Display Size) × 3 (Instruction Type) × 2 (Position) × 2 (Gender) repeated measures analysis of variance (RM-ANOVA). The Position variable refers to the order in which a given user worked on each level of the Display variable, and allowed us to explore ordering as well as skill transfer effects across Displays. We analyzed Instruction Type, Position, and Gender as between-subjects factors and Display Size as a within-subject factor.

Experiment 1: Average Percentage of Correct Responses

Fig. 5.   Main effects of Strategy, with users performing significantly better with Egocentric Instructions than Exocentric ones. Also, results suggest that users with Unbiased Instructions perform with exocentric strategies when using the Small Display, and with egocentric strategies when using the Large Display. Error bars represent standard error.

Overall, we found a significant effect of Instruction Type ($F(2, 37) = 3.866$, $p = .030$; see Figure 5). Paired comparisons showed a significant difference between the egocentric and the exocentric instruction sets ($p = .01$), with users getting a higher percentage of questions correct with egocentric instructions than the exocentric ones (66.5% vs. 47.2%, respectively).

We conducted planned contrasts to see if users who were explicitly instructed to use a given strategy performed any differently from users who implicitly chose a strategy due to the Display Size. We found no significant differences between users in the exocentric condition and the unbiased small display condition, which was assumed to elicit an exocentric strategy ($t(40) = .079$, $p = .9371$). Similarly, we found no significant differences between users in the egocentric condition and the unbiased large display condition, assumed to elicit an egocentric strategy ($t(40) = 0.953$, $p = .3463$). Rather than pooling Display Size in the exocentric and egocentric conditions, we also conducted tests comparing performance on the small display in the exocentric condition to the small display in the unbiased condition, as well as the large display in the egocentric condition to the large display in the unbiased condition. In both cases, there were no significant differences. These results, seen in Figure 5, replicate findings from the initial experiments as well as provide support for our hypothesis that large displays provide a greater sense of presence and bias users into using egocentric strategies.

5.3.2 *Preference Data.*   We gathered preference data from the participants at the conclusion of the experiment. The merged preference data for all three Instruction Type conditions were not significantly in favor of the large display. We explored whether or not users in the different Instruction Types viewed the value of the displays differently. We found in paired comparisons that users with the egocentric instructions and the unbiased instructions preferred the large display significantly more than users given the exocentric instructions for 'Ease

of Seeing' (p = .034 and p = .046, respectively) and marginally significantly more in 'Confidence in Rotation' (p = .064 and p = .077, respectively). However, we did not see any significant differences in 'Overall Preference' across Instruction Types, suggesting that effects were probably not driven entirely by display characteristics and subjective preference. In general, these satisfaction ratings complement the performance results nicely.

5.4 Summary

Although we were not able to directly measure choice of cognitive strategy in our pilot experiments, results from this experiment indicate that users performed better when they used an egocentric strategy than when they used an exocentric one. Also, simple instructions and training were indeed sufficient to bias users into adopting one or the other of the strategies. In the absence of an explicit strategy, users seem to have chosen an exocentric one when working on the small display and the much more efficient egocentric one when working on the large display. Results from this experiment suggest that, given a constant visual angle, the size of a display implicitly affects choice of cognitive strategy and hence performance in spatial orientation tasks. While we assert that this strategy choice is due to large displays offering a greater level of immersion, explicitly demonstrating the causality remains future work.

In the next experiment, we provide additional support and insight into this explanation. If the explanation is correct, and the cause of the observed performance benefits is the implicit choice of an egocentric strategy, we would expect not to see benefits in tasks for which egocentric strategies do not help.

6. EXPERIMENT 2: LARGE DISPLAYS DO NOT AID EXOCENTRIC TASKS

While Guilford [1972] considered a single spatial orientation factor in his work, other researchers (e.g. Lohman [1979]) have identified three related spatial ability factors: spatial egocentrism, the ability of the observer to imagine their body in a different position so that they can figure out how a stimulus array will appear from another perspective; spatial relations, the ability to identify certain objects when seen from different positions; and visualization, the ability to form a mental image of something that is not visible.

The Guilford-Zimmerman Spatial Orientation test used in the first experiment allowed the user to use either spatial egocentrism or exocentric spatial relations strategies to perform the task. It was the choice of these strategies, biased either by prior instructions or by the size of the display that accounted for the observed performance differences. In this experiment, we picked tasks that did not seem like they would benefit from doing the task with a spatially egocentric strategy. Thus,

> *Hypothesis 2: Large displays bias users into using egocentric strategies and do not increase performance on 'intrinsically exocentric' tasks for which egocentric strategies are not useful.*

6.1 Participants

Twenty-four (12 female) college students from the Pittsburgh area, who did not participate in the previous experiments, participated in this one. As before, we screened users to have normal or corrected-to-normal eyesight. The average age of users was 24.1 (24.4 for males, 23.8 for females), ranging from 18 to 44 years of age. The experiment took about an hour and users were paid for their participation.

6.2 Procedure

We used the same hardware setup as in the previous experiments. The tasks used in this experiment were selected because they are object-centric problems and have been described as inherently exocentric tasks that would not benefit from having the user imagine their body within the problem space (i.e., they would not benefit from the participants using an egocentric strategy). The first two tasks, the Card test and the Cube test, are subtests S-1 and S-2 of the ETS Kit of Factor-Referenced Cognitive Tests [Ekstrom et al. 1976]. The tests are inspired by Thurstone's cards and cubes [Thurstone and Thurstone 1941]. The third task, the Shepard-Metzler test [Shepard and Metzler 1971], is a task commonly used to study mental rotation. We used a subset of questions from this test.

Before beginning the tasks, subjects filled out a background questionnaire and adjusted themselves in the chair so that their eyes were as close to the appropriate point in the room that ensured equivalent visual angles between displays. Subjects then did each of the Card Test, the Cube Test, and the Shepard-Metzler Test, in that order. This experiment was a within-subjects design, with each subject performing each of the tasks in both the Large Display and Small Display conditions in an order that was counterbalanced between subjects. Finally, they completed a preference survey.

6.2.1 *Card Test.*   In each question of this test, the user saw two cards, each with the image of an irregular shape (see Figure 6). The two cards showed either the same shape or mirror images of the shape, rotated to different degrees. The user's task was to mentally rotate the cards in the plane and determine if they represented the same shape or if they were mirror images of each other.

The original paper-based test presented a single base image to which eight other images were compared. Each section of the test was printed on a single page with 10 such rows of questions, for a total of 80 questions. In the computer-based version of this test, we showed each pair of cards one pair at a time, advancing to the next pair only when the user responded to the question. The left card in each pair corresponded to the base shape in the paper-based test. Users had 3 minutes to complete each set of 80 questions.

6.2.2 *Cube Test.*   In each question of this test, the user saw two cards, each with the drawing of a cube containing different characters in different orientations on each face (see Figure 6). Users were told that no character appeared on more than one face of a given cube. The user's task was to mentally rotate the
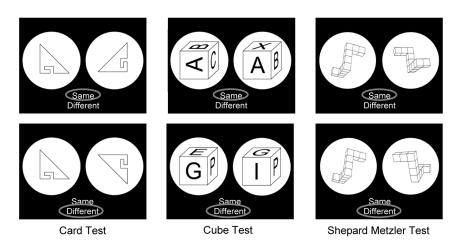
Fig. 6. Exocentric tests that do not benefit from egocentric strategies, or user imagining their bodies within the problem space. In each, the user has to mentally rotate images to determine if they can be of the same object.

cubes and determine if the drawings could have represented the same physical cube, or if they were definitely different cubes.
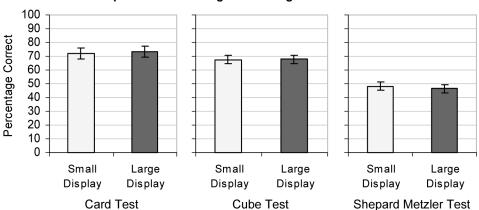
Similar to the Card test, the paper-based test presented each set of 21 distinct pairs simultaneously on a single page. In the computer-based version, we showed each pair one at a time, again advancing only when the user had provided an answer. Users had 3 minutes to complete each set of 21 questions.

6.2.3 *Shepard-Metzler Test.* This test is similar to the Card test except that the mental rotation task is three-dimensional. Each question presents two drawings of objects in space (see Figure 6). Each object consists of 10 solid cubes attached face-to-face to form a rigid arm-like structure. Users had to mentally rotate the objects in space in order to determine if they were the same object, or if they were different. Once they indicated their answer, the system advanced to the next question.

The original Shepard-Metzler stimuli of 70 line drawings consisted of 10 different objects in 7 positions of rotation about a vertical axis. These 7 positions permit the construction of at least two unique pairs at each angular difference in orientation from 0 to 180 degrees, in 20-degree increments. In this experiment, we created two equivalent subsets of the test, each with 60 questions: 6 objects $\times$ 5 angles (20, 60, 100, 140 and 180 degrees) $\times$ 2 answers (same or different object). We presented each pair to users one at a time. Users had no time limit for this task, but were reminded to perform the questions as quickly and accurately as possible.

6.3 Results

As before, we first present performance on the spatial orientation task, and then we examine the preference data.

Experiment 2: Average Percentage Correct for each Test

Fig. 7.   Users performed no differently on any of the tasks whether using the Small or the Large Display. Egocentric strategies do not help on exocentric tasks. Error bars represent standard error.

6.3.1 *Exocentric Task Performance.*   Since we did not expect effects across experiments, we analyzed each of the tests independently. We modeled the data for each of the three tasks at the summary level, analyzing the percentage of correct responses (number correct/number attempted) for each test with a 2 (Display Size) $\times$ 2 (Position) $\times$ 2 (Gender) repeated measures analysis of variance (RM-ANOVA). As before, Position refers to the order in which users saw each level of the Display factor. We saw similar results when we used the absolute number of correct answers as the dependent measure. We analyzed Gender and Position as between-subjects factors and Display Size as a within-subjects factor.

We saw no effects of Display Size in each of the three tests, with no significant difference in percentage of correct responses for the Card test ($F(1, 19) = 1.473$, $p = .240$), the Cube test ($F(1, 19) = 0.012$, $p = .914$), or the Shepard-Metzler test ($F(1, 19) = 0.5108$, $p = .475$). The effect sizes for these tests ranged from negligible to small/medium ($d = .56$, $d = .05$, and $d = .32$, respectively). Hence we are most confident in the results for the Cube test and least confident in those for the Card test. However, given the converging results of all three tests, we believe that the egocentric benefits do not impact these exocentric tasks. These results can be seen in Figure 7. Likewise, none of the other main effects or interactions was significant for this dependent measure.

When we compared the average time spent per question on each of the tests, we found no significant interactions with the display manipulation. One point worth noting is that when we conducted an analysis at trial level, similar to that performed in the original Shepard and Metzler [1971] experiments, we found comparable results. We found significant effects ($F(1, 2267) = 32.704$, $p < .001$) suggesting that the larger the angle of mental rotation required to align the two objects, the longer it took users to decide whether the objects were the same or if they were different. In fact, the relationship between angle of rotation and time spent on the question was a linear trend, as predicted.

6.3.2 *Preference Data.* Overall we found no significant differences in preference when users were asked to rate the two displays on a 5-point Likert scale of 1 = "Strongly Disagree" to 5 = "Strongly Agree". The questions were 'information on this display was easy to see' (M = 4.33 vs. M = 4.13 for small vs. large display), 'the task was easy to do on this display' (M = 3.79 vs. M = 3.70 for small vs. large display), and 'overall I liked this display' (M = 4.13 vs. 3.79 for small vs. large display). This corresponds well with performance data.

6.4 Summary

Even though there is evidence that the tests used in this experiment utilize similar cognitive abilities as the Guilford-Zimmerman task, namely spatial orientation and mental rotation, we asserted that users would not benefit from imagining their bodies within the problem space due to the object-centric nature of stimuli. There was no reason to believe that imagining their body within the problem space would help with these tasks. Results indeed showed that users did not experience the same benefits on these exocentric tasks that they did on the Guilford-Zimmerman task. In fact, users performed just as well when they worked on the small display as on the large display. This finding provides additional support to the explanation that performance benefits were due to an increased sense of presence that biased users into egocentric strategies, strategies that were not useful for intrinsically exocentric tasks. It also implies that we must be very careful in applying the finding as large display benefits only apply to tasks that can be performed more effectively using egocentric strategies.

It was initially advantageous to use well-validated and established psychology tests in order to understand the particular psychophysical phenomena in which we were interested. Although effects were easy to interpret, these tests had several shortcomings: (1) they were designed to isolate and study very controlled spatial abilities and did not take into account tasks in which compound abilities would be used; (2) the stimuli were often contrived two-dimensional, black and white images; and (3) they were static multiple choice tests that did not require the user to interact with the virtual environment.

In the following experiments, we extend the work by applying findings to more ecologically valid tasks. We incrementally increase the complexity of spatial abilities used in order to see if the current effects continue to be robust. We also use fairly rich dynamic three-dimensional virtual environments and incrementally increase the complexity of these environments by adding cues such as distinct landmarks and textures in order to see how the effects hold up in the presence of other cues. Finally, we test for the reliability of the large display effects when the user actively interacts with the virtual environment. Interactivity could potentially immerse the user within a virtual environment and cause them to perform better, hence negating some of the benefits afforded by large displays.

7. EXPERIMENT 3: LARGE DISPLAYS AID MAP FORMATION AND MEMORY

Results discussed thus far show that information presented on physically large displays provides a greater level of immersion and allows users to perform

certain tasks more effectively than on smaller desktop displays, even when information is viewed at equivalent visual angles.

In separate work, we have also shown that users perform 3D navigation tasks requiring path integration more efficiently on large displays than on smaller ones, even when identical scenes were viewed at equivalent visual angles [Tan et al. 2004]. In those experiments, we have further shown that the distraction imposed by active navigation control using a joystick may outweigh any additional cues it might have provided, at least for the set of tasks we tested. However, effects induced by interactivity seem to be independent of those induced by display size. Our follow-up investigation showed that locomotion errors were small and that our results could mainly be attributed to wayfinding errors.

In this experiment, we extend these results to include a mental map formation and memory task. In this task, the user explores a virtual world in order to build a cognitive map of the environment. Using this cognitive map, the user then navigates to several specified targets as quickly as they can. Users who build and remember better cognitive maps should be able to navigate to the targets with shorter distances and in less time.

There exists a vast body of work on general principles in 3D navigation. Thorndyke and Hayes-Roth [1982], as well as many others (e.g. Ruddle et al. [1999]; Waller et al. [1998]), have studied the differences in spatial knowledge acquired from maps and exploration. Darken and others have explored cognitive and design principles as they apply to large virtual worlds [Darken and Sibert 1996]. All this work recognizes that 3D navigation is a complex cognitive task requiring the use of a series of interrelated spatial abilities. We believe that benefits of large displays for simple spatial tasks extend to more complex ones, hence,

> *Hypothesis 3a: Users perform better in mental map formation and memory tasks when using physically large displays due to the increased likelihood that they adopt egocentric strategies.*

In separate work, some researchers have found that the acquisition of spatial knowledge is facilitated by active navigation control (e.g. Cutmore et al. [2000]; Philbeck et al. [2001]). These researchers claim that proprioceptive cues provided by the input devices as well as cognitive benefits of decision-making immerse users more within the virtual environments and aid in encoding mental representations of the environments. Others however, have reported opposite results, showing that active control hurts performance in various navigation tasks (e.g. Booth et al. [2000]). Flach [1990] argues that the different results could be due to the tradeoffs imposed by control of attention, kinds of information available, sensitivity to information, as well as activities involved.

We decided to explore both how level of interactivity in the virtual environment affects mental map formation and memory in our 3D navigation task, as well as how it interacts with effects caused by varying the physical size of displays. While prior literature provides evidence of active control helping in some situations and hurting in others, we expected users to perform better when they had interactive control using the joystick due to the additional cues afforded by
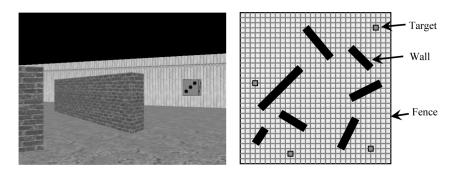
Fig. 8.   (left) First person view of the world, including walls, target, and fence (right). Map view of an example world. The user never saw this view.

the physical manipulation. Therefore,

> *Hypothesis 3b: Users perform better in the path integration task when they are interactively moving themselves through the virtual environment.*

Finally, we expected the benefits of the large display to be robust against other factors that could potentially provide a similar heightened sense of presence. Specifically,

> *Hypothesis 3c: The effects induced by physical display size are independent of those induced by interactivity.*

7.1 Participants

Sixteen (8 female) intermediate to experienced computer users from the Greater Puget Sound area participated in the experiment. We screened users to be nongamers who played less than 3 hours of video games per week. We also screened users to be fluent in English and to have normal or corrected-to-normal eyesight. The average age of users was 36.0 (33.7 for males, 38.3 for females), ranging from 19 to 47 years of age. The experiment took about an hour and a half and users were given software gratuities for their participation.

7.2 Procedure

We created five different 3D virtual worlds using Touchdown Entertainment's Jupiter game development platform. Each of these worlds was a square room with edges 30 meters long. The room was bounded by a fence to prevent users from wandering outside of it (see Figure 8). Seven walls were randomly distributed throughout the environment. To ensure a well-distributed pattern of walls, we ensured that: (1) the average length of walls, approximately 7 meters, was comparable across the worlds, and (2) that each quadrant of the world had a roughly equivalent number and length of wall segments. We then distributed four red target cubes, one in each quadrant of the world. Each cube was uniquely marked and could be identified by the number of dots (one to four) found on each of its faces. Within this world users had basic joystick controls. Pushing the stick forward and backward, moved them forward and backward; pushing

Fig. 9.   The joystick used (left); User working on the small display (center) and the large display (right).

the stick left or right turned them left or right. They moved at a maximum speed of 2.5 meters per second, and turned at a maximum rate of 30 degrees per second.

We used a mental map formation and memory task to test how each of the manipulations affected the way users performed in various 3D virtual environments. We broke the task into two phases: the learning phase, and the recall phase. In the learning phase, we gave users 4 minutes to explore the world and learn both the structure of the environment as well as the location of the various target cubes within the world. In the recall phase, we placed users in random locations within the world and had them move to specified targets as quickly as possible. These random locations were chosen such that the optimal path to the specified target was always 20 meters long. Users were asked to find each of the target cubes twice, for a total of eight trials per world. Note that the environments did not contain any distinct landmarks or textures (see Figure 8). The only way to remember the location of targets was to build a mental map using the structure of walls within the environment.

Dependent measures in this experiment included: (a) the distance moved from the start-point to the target in the recall phase; (b) the time required for the user to find each of the targets.

This was a 2 (Display Size: Small vs. Large) × 2 (Interactivity: Passive Viewing vs. Active Joystick) within-subjects design. In the Active Joystick condition, users utilized the joystick to move themselves through the environment as they explored it in the learning phase (see Figure 9). In the Passive Viewing condition, users had no control of their movement in the learning phase. Instead, they viewed a movie of themselves moving through the environment. In pilot tests, we used the output from one user's Active Joystick condition as the stimulus for another's Passive Viewing condition. However, we found that most users moved themselves in somewhat unpredictable motions through the environment. This either caused an unreasonably high level of motion sickness in viewers, or was so jerky as to be ineffective in helping to learn the environment. Hence for the Passive Viewing condition we scripted a smooth path designed to explore the environment by moving in between every pair of walls at least twice in a systematic manner.

Prior to beginning the test, users completed a background questionnaire as well as a spatial ability test. We used the Paper Folding test (VZ-2) from the

Kit of Factor-Referenced Cognitive Tests [Ekstrom et al. 1976]. This test is a well-validated spatial orientation test that is commonly used to indicate general spatial ability. Users were then given detailed instructions and performed the task in the tutorial world. Following this, users performed the task in all four conditions, each in one of the four different environments. The conditions and the specific worlds were balanced across users. Finally, users filled out a preference questionnaire.

7.3 Results

We used a mixed model analysis of variance (ANOVA) in which Display Size and Interactivity were repeated and Gender was treated as a between-subjects factor. We included all 2-way and 3-way interactions in the analysis. Since this was a completely within-subjects design, observations were not independent and we modeled User as a random effect nested within Gender. We originally included two covariates in the model: distance moved in the learning phase, and the Spatial Abilities score. However, we removed these from the final analyses because they were not significant. Distance moved in the learning phase was not significantly different in any of the conditions, and Spatial Abilities did not interact with any of the manipulations. The estimates and significance levels of the main factors of interest did not change in any significant fashion and the overall model fit was improved.

In this experiment, we found a significant effect of Display Size ($F(1, 487) = 26.745$, $p < .001$), with the Large Display resulting in users moving shorter distances to find the targets than the Small one (35.31 vs. 39.93 meters, respectively). We also observed a significant effect of Interactivity ($F(1, 487) = 14.219$, $p < .001$), with trials in the Active Joystick condition demonstrating shorter distances moved than trials in the Passive Viewing condition (20.94 vs. 24.30 meters, respectively). Interacting with the environment seemed to aid users with the map formation and memory task. We saw no interaction between Display Size and Interactivity ($F(1, 487) = 0.909$, $p = .341$), suggesting that the manipulations were independent of one another. These results can be seen in Figure 10. Finally, we saw a main effect of Gender ($F(1, 487) = 9.119$, $p = .003$), with males performing better than females (32.25 vs. 42.97, respectively). We saw no interactions between Gender and any of the manipulations.

The results for time required to find each of the targets matched findings using the distance moved metric. We found significant effects of Display Size ($F(1, 487) = 71.179$, $p < .001$), Interactivity ($F(1, 487) = 38.026$, $p < .001$), and Gender ($F(1, 487) = 5.259$, $p = .022$). There were no other significant effects or interactions. This is not surprising, as many subjects moved around at close to the top speeds even when they did not immediately know their way around the environment.

7.4 Summary

This experiment adds further validity to our previous findings. In this experiment, we continued to see benefits of using large displays even with a fairly complex task requiring the use of numerous spatial skills. Like other tasks

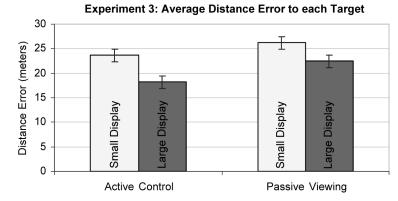**Experiment 3: Average Distance Error to each Target**

Fig. 10. Main effects of Interactivity and Display Size. In this experiment, users benefited from Active Control. There were again no interactions between manipulations. Error bars represent standard error.

that benefit from the use of large displays, the map formation and memory task benefited from having users adopt an egocentric frame of reference while navigating.

In this experiment, we found that active navigation control helped users learn and remember environments more effectively. In fact, they performed about 10% better when they controlled their movement than when they watched a video of themselves moving through the environment. Interestingly, and importantly, effects of interactivity were still independent of effects induced by display size.

One shortcoming of this experiment is that the virtual environments used were still fairly sterile and controlled. They did not contain distinct landmarks or textures, which would be expected to exist in a more ecologically valid environment. We did this so that we could better understand the nature of the task and basic results before moving into a more complex environment in which other factors could contribute to effects observed. We conducted the next experiment to explore how robust these effects were in the presence of a multitude of additional cues found in more typical virtual environments.

8. EXPERIMENT 4: ECOLOGICAL VALIDITY OF RESULTS

This experiment extends previous results by testing the effects of display size in a much more ecologically valid environment. Thus,

> *Hypothesis 4: Even in an environment crafted with cues such as distinct landmarks and rich textures to be realistic and memorable, users perform better in mental map formation and memory tasks when using physically large displays due to the increased likelihood that they adopt egocentric strategies.*

8.1 Participants

Sixteen (8 female) college students from the Pittsburgh area who were intermediate to experienced computer users participated in the study. We screened

Fig. 11.   First person view of the world, which contains distinct landmarks and rich textures in Experiment 4. The target is the flag seen in the lower right of the screen.

users to be nongamers who played less than 3 hours of video games per week. We also screened users to be fluent in English and to have normal or corrected-to-normal eyesight. The average age of users was 23.9 (24.6 for males, 23.1 for females), ranging from 18 to 31 years of age. The experiment took about an hour and a half and users were paid for their participation.

8.2 Procedure

We used the same mental map formation and memory task as in Experiment 3. In the learning phase, users were allowed to explore a virtual environment for 4 minutes. In the recall phase, they had to locate specific targets from random locations within the world. These locations were randomly distributed to be distances of 50 meters away from the targets.

We used an off-the-shelf copy of Unreal Tournament 2003 (Epic Games) for this experiment. Unreal Tournament is a first-person shooter and can be considered to utilize a state-of-the-art rendering engine and virtual environments. In fact, virtual environments in this game are specifically crafted to be realistic, or immersive, and memorable for players (see Figure 11). Unreal Tournament comes with a game-mode called 'capture the flag.' Each of the worlds in this game has two team flags. In order to score, one team must touch the enemy flag and return it to their home base. We used these flags as targets to find within the environment.

Unreal Tournament comes with development tools for editing maps as well as for scripting simple behaviors within the worlds. We made two modifications to the game in order to run the experiment. First, we instrumented the game so that we could log the dependent measures: (a) the distance moved from the start-point to the target in the recall phase; (b) the time required for the user to find each of the targets. Second, we had initially left several computer enemies in the game to serve as further distraction while the user performed the mental map formation and memory task. However, in pilot tests, users got so carried

away chasing and shooting enemies that they forgot all about their main task. As such, we removed all enemy characters as well as weapon pickups from the worlds in the actual tests.

We chose worlds from the standard set of worlds that ship with the game as well as from upgrade packages created by gamers and distributed on various websites. Through pilot tests, we selected five of these worlds from an initial pool of twelve, such that we had one small tutorial world, two Easy Worlds, and two Difficult Worlds. The Easy Worlds both covered about 1000 square meters and the Difficult Worlds covered a little more than twice that amount of space. Additionally, the Difficult Worlds were much harder to learn and navigate due to the complexity of structures and cues within the environment. For example, one such world had a maze of underground caverns and tunnels to navigate. Pilot tests suggested that each pair of worlds, the Easy Worlds and the Difficult Worlds, were of roughly similar difficulty within our task.

We eliminated the Interactivity manipulation from the previous experiment, and hence this experiment was a 2 Display Size (small vs. large) $\times$ 2 Difficulty (easy vs. hard) within-subjects design. The orders of Display Size and Difficulty were independently balanced between users.

8.3 Results

We used the same analysis model as in the previous two experiments, replacing the Interactivity manipulation with the Difficulty one in this experiment. While we saw dampened effect sizes from the previous experiment, possibly due to the dilution caused by the addition of cues within the environments, we observed similar findings. We found a significant effect of Display Size ($F(1, 604) = 11.900$, $p < .001$), with the large display resulting in users detouring by shorter distances to find the targets than the small one (14.71 vs. 16.35 meters, respectively). We also observed a significant effect of Difficulty ($F(3, 604) = 108.996$, $p < .001$), with trials in the easy worlds demonstrating shorter distances moved than trials in the difficult worlds. We saw a significant interaction between Display Size and Difficulty ($F(3, 604) = 4.041$, $p = .007$). The large display seemed to be more helpful with trials in the difficult worlds. Finally, we saw a main effect of Gender ($F(1, 604) = 6.699$, $p = .010$), with males performing better than females (12.94 vs. 18.14 meters, respectively). We saw no interactions between Gender and any of the manipulations.

Again, the results for time required to find each of the targets matched findings using the distance moved metric, with a significant effect of Display Size ($F(1, 604) = 4.281$, $p = .039$), Difficulty ($F(3, 604) = 294.510$, $p < .001$), and Gender ($F(1, 604) = 7.319$, $p = .007$), but no other main effects or interactions.

8.4 Summary

This experiment shows that benefits of large displays are independent of cues that may be used in real-world virtual environments to increase immersion and memorability, such as distinct landmarks and rich textures. This is an important property if we are to apply the summary of results to useful real-world tasks, such as training and simulation, or games and entertainment.

Also, it implies that we can continue to exploit the benefits of large displays even in the presence of other techniques that induce performance increases.

9. GENERAL DISCUSSION OF EXPERIMENTAL RESULTS

The series of experiments described in this article demonstrate that physical display size is an important factor to consider when designing display systems. Results suggest that physically large displays, even at identical visual angles as small displays, immerse users and bias them into adopting egocentric strategies. These strategies increase performance on spatial tasks such as 3D navigation as well as mental map formation and memory, which can be represented using egocentric coordinate systems. Furthermore, the effects caused by physically large displays seem to be independent of other factors that may induce immersion or increase performance. For example, even though interactivity and mental aids such as distinct landmarks and rich textures within virtual worlds increase task performance on the tasks tested, they did not affect the benefits that large displays offer to users.

In fact, with very little effort on the part of the designer, the system builder, or the user, large displays offer the potential to improve performance on a fairly broad range of tasks. Also, because effects are independent of other aids tested, large displays continue to offer improvements even in the presence of other performance aids.

It could be argued that the magnitude of effects was not amazingly large, and that 10% to 26% increases are not enough to warrant the additional cost and physical space that large displays require. However, given that the theoretical information content shown on the small and the large displays were the same, and hence that the retinal images created when viewing one display or the other were the same, it is interesting that these results exist at all.

Furthermore, it should be noted that performance gains even of this magnitude could be important in the domains for which we think these results are most useful, namely games and entertainment, as well as training and simulation. Games and entertainment is a large market that continues to grow, and that could benefit significantly from even a small portion of the demographic preferring and upgrading to large displays. In training and simulation, any small increase in performance could potentially lead to fairly large implications. For example, imagine firefighters who could navigate to targets 10% quicker or could better find alternate routes when they become obstructed because they trained on large displays. Before we can build such applications, we must explore how strategies learned on large displays transfer to the real world, especially since we did not see training transfer results between displays in any of our experiments. This remains future work.

The behavioral effect and choice of different strategy depending on the physical size of the display is perhaps more interesting than the raw magnitude of performance increases. The magnitude of the effect is heavily dependent on the particular task as well as the surrounding context in which the task is performed. However, the behavioral effect can be attributed to a much more fundamental cognitive mechanism, which may form an important component

of the way we perceive and interact with the world around us. Even though we were not able to directly measure causality, the experiments as a group provide strong support for our hypothesis that the performance increase on large displays can be attributed to the choice of cognitive strategy. We hypothesize that this strategy switch was induced in part by a higher level of immersion when using the large display. While this has been postulated in previous literature [e.g. Bystrom et al. 1999], large displays have also been shown to mediate other psychological factors such as emotional arousal and attention [Reeves et al. 1999]. Validating exactly why users change their cognitive strategy remains future work.

It should also be noted how robust the results were to the types of tasks tested, as well as to the demographic for which this applied. Because the experiments were performed both with college students from Carnegie Mellon University as well as with a wide range of people recruited from the general population in the Greater Puget Sound area, we can say with relatively high degree of confidence that the results are representative of a large portion of the population.

Informal observations across the demographic yielded other interesting design considerations. For example, people with bifocals usually preferred reading and performing the tasks on the large display. This was because they were much more comfortable working on surfaces that were further away. Depending on demographic, users would compare the large display to a movie screen or a classroom board, but most indicated that they were more engrossed by the large display. Unfortunately, it was difficult to get users to articulate the level of immersion or the actual strategy they used to perform tasks. In fact, in pilot experiments as well as in the actual experiments, we tried various methods including multiple choice and ranking questions, magnitude questions using Likert scales, subjective open-answer questions, and informal interviews. While preference ratings generally matched performance results, none of these methods was effective in deriving definitive responses or insights regarding strategy used. Instead, we had to resort to carefully designing the experiments such that we ended up with a series of performance results suggesting that the strategy hypothesis is the most likely explanation for the effects observed.

As a final note, although we observed effects of gender and spatial ability across many of these tasks, we did not pursue these further. These effects have been fairly well documented in the literature and were not the focus of our experiments. While Czerwinski et al. [2002] suggest that females benefit significantly more than males in 3D navigation tasks using displays with wide fields of view, we saw no such effect for Display Size in our studies. We found no interactions between any of the manipulations and these factors, indicating that nothing surprising was happening with these effects. We have found no evidence suggesting that physical display size aids any part of the demographic more or less so than any other group.

10. FUTURE WORK

Although we did not intentionally calibrate the absolute size of the images in any of the experiments, images shown on the large display were close to being

life-sized. This might be an interesting point in the size-performance curve as it is usually assumed to represent the optimal size at which users will be immersed (e.g. Brooks [1999]). However, more work is required to determine the shape of the curve as one increases display size from a traditional desktop display to a large wall-sized display and then beyond. It would be interesting to see if the strategy change is an abrupt shift that happens when a certain size is achieved or if it is more continuous across a series of sizes. Also, it would be interesting to see what happens when images get larger than life. This would allow us to gain a deeper understanding of display size and how it relates to immersion and presence.

Another potentially interesting realm of study has to do with the factors that best allow us to perceive physical size. There are numerous factors such as optical accommodation and convergence, stereo vision, parallax, and environmental context, but we do not have a clear understanding of how each of these contributes to the effects and how they interact with each other. We believe that this could form an entire research agenda, and would add not only to our understanding of large displays, but of the human visual and perceptual systems in general.

Building upon an understanding of what it is that allows us to perceive size, we believe that it is also important for us to completely understand what it is about that size that causes us to become more immersed and to adopt different strategies when performing spatial tasks. For example, if these results could be partially attributed to the novelty of the large displays, then results would be a little less useful theoretically, but it would be very interesting to find out that they were due to certain fundamental biases in our neural circuitry. Again, this remains future work.

Before we fully understand the design principles derived from these experiments within real world scenarios, there are a few other areas to consider. For example, we must look beyond behavioral responses and performance results, and understand how large displays affect other things such as simulator sickness. In our experiments, we saw no indication that large displays would cause any more or less illness, but we cannot draw any conclusions because almost no one got sick in the experiments.

Also, we must fully examine the interaction of other display characteristics, such as field of view or resolution. Maintaining constant visual angles was merely a means to isolate and study physical display size as an interesting characteristic. We do not propose that large displays should be intentionally used at equivalent visual angles to their smaller counterparts. Instead, we should clearly understand the interactions between display size and these other factors so that we can design display systems that make optimal use of large displays.

Finally, we believe that we must explore training transfer between different types of displays and between the virtual world and the real one. None of our experiments showed any ordering effects, suggesting that the strategy change was a rather ephemeral one, changing quickly, depending on which display the user was currently using. However, given the length of our tasks, we would be hesitant to draw conclusions about any longer term training transfer effects; this will have to be studied in more detail.

REFERENCES

ARTHUR, K. W. 2000. Effects of field of view on performance with head-mounted displays. *Doctoral Dissertation, University of North Carolina*, Chapel Hill, NC.

BAUDISCH, P., GOOD, N., BELLOTI, V., AND SCHRAEDLEY, P. 2002. Keeping things in context: A comparitive evaluation of focus plus context screens, overviews, and zooming. In *Proceedings of the CHI 2002 Conference on Human Factors in Computing Systems*, 259–266.

BOOTH, K., FISHER, B., PAGE, S., WARE, C., AND WIDEN, S. 2000. Wayfinding in a virtual environment. *Graphics Interface 2000*.

BROOKS, F. P. 1999. What's real about virtual reality? *IEEE Computer Graphics and Applications*, 19, 6, 16–27.

BUXTON, W. 2001. Less is more (More or less). In *The Invisible Future: The Seamless Integration of Technology in Everyday Life*. P. Denning, Ed. McGraw Hill: New York, NY, 145–179.

BYSTROM, K. E., BARFIELD, W., AND HENDRIX, C. 1999. A conceptual model of the sense of presence in virtual environments. *Presence: Teleoperators and Virtual Environments 8*, 2, 241–244.

CARPENTER, M. AND PROFFITT, D. 2001. Comparing viewer and array mental rotations in different planes. *Memory & Cognition 29*, 441–448.

CHAPANIS, A. AND SCARPA, L. C. 1967. Readability of dials at difference distances with constant viewing angle. *Human Factors 9*, 5, 419–426.

CHILDS, I. 1988. HDTV-putting you in the picture. *IEE Rev. 34*, 7, 261–265.

CHOU, P., GRUTESER, M., LAI, J., LEVAS, A., MCFADDIN, S., PINHANEZ, C., VIVEROS, M., WONG, D., AND YOSHIHAMA, S. 2001. BlueSpace: Creating a personalized and context-aware workspace. *IBM Tech. Rep. RC22281*.

CUTMORE, T. R. H., HINE, T. J., MABERLY, K. J., LANGFORD, N. M., AND HAWGOOD, G. 2000. Cognitive and gender factors influencing navigation in a virtual environment. *Int. J. Hum. Comput. Studies 53*, 223–249.

CZERWINSKI, M., TAN, D. S., AND ROBERTSON, G. G. 2002. Women take a wider view. In *Proceedings of the CHI 2002 Conference on Human Factors in Computing Systems*, 195–202.

DARKEN, R. AND SIBERT, J. 1996. Navigating in large virtual worlds. *Int. J. Hum.-Comput. Interaction 8*, 1, 49–72.

DUDFIELD, H. J., MACKLIN, C., FEARNLEY, R., SIMPSON, A., AND HALL, P. 2001. Big is better? Human factors issues of large screen displays with military command teams. In *Proceedings of People in Control 2001*, 304–309.

EKSTROM, R. B., FRENCH, J. W., HARMAN, H., AND DERMEN, D. 1976. Kit of factor-referenced cognitive tests. In *Educational Testing Service*: Princeton, NJ.

ELROD, S., BWEE, R., GOLD, R., GOLDBERG, D., HALASZ, F., JANSSEN, W., LEE, D., MCCAU K., PEDERSEN, E., PIER, K., TANG, J., AND WELCH, B. 1992. Liveboard: A large interactive display supporting group meetings, presentations and remote collaboration. In *Proceedings of the CHI 1992 Conference on Human Factors in Computing Systems*, 599–607.

EPIC GAMES. Unreal Tournament. http://www.unrealtournament.com

FLACH, J. 1990. Control with an eye for perception: Precursors to an active psychophysics. *Ecol. Psych. 2*, 83–111.

GUILFORD, J. P. 1972. Thurstone's primary mental abilities and structure-of-intellect abilities. *Psyc. Bull. 77*, 2, 129–143.

GUILFORD, J. P. AND ZIMMERMAN, W. S. 1948. The Guilford-Zimmerman aptitude survey. *J. Appl. Psych. 32*, 24–34.

GUIMBRETIÈRE, F. 2002. Fluid interaction for high resolution wall-size displays. Doctoral Dissertation, Stanford University, Stanford, CA.

INFIELD, S. E.   1991.   An investigation into the relationship between navigation skill and spatial abilities. *Doctoral Dissertation, University of Washington*, Seattle, WA. Dissertation Abstracts International, 52(5-B), 2800.

ISHII, H. AND ULLMER, B.   1997.   Tangible bits: Towards seamless interfaces between people, bits and atoms. In *Proceedings of the CHI 1997 Conference on Human Factors in Computing Systems*, 234–241.

LIN, J. J., DUH, H. B. L., PARKER, D. E., ABI-RACHED, H., AND FURNESS, T. A.   2002.   Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment. In *Proceedings of IEEE Virtual Reality Conference 2002*, 164–171.

LOHMAN, D.   1979.   Spatial ability: A review and reanalysis of the correlational literature. *Tech. Rep., N.8. Stanford University*, Aptitude Research Project, School of Education.

PATRICK, E., COSGROVE, D., SLAVKOVIC, A., RODE, J. A., VEWATTI, T., AND CHISELKO, G.   2000.   Using a large projection screen as an alternative to head-mounted displays for virtual environments. In *Proceedings of the CHI 2000 Conference on Human Factors in Computing Systems*, 478–485.

PHILBECK, J. W., KLATZKY, R. L., BEHRMANN, M., LOOMIS, J. M., AND GOODRIDGE, J.   2001.   Active control of locomotion facilitates nonvisual navigation. *J. Exper. Psych. Human Perception and Performance 27*, 141–153.

PROTHERO, J. D. AND HOFFMAN, H. D.   1995.   Widening the field of view increases the sense of presence within immersive virtual environments. *Human Interface Technology Laboratory Tech. Rep., University of Washington*, Seattle, WA, R-95-4.

RASKAR, R., WETCH, G., CUTTS, M., LAKE, A., STESIN, L., AND FUCHS, H.   1998.   The office of the future: A unified approach to image-based modeling and spatially immersive displays. In *Proceedings of SIGGRAPH 1998 International Conference on Computer Graphics and Interactive Techniques*, 179–188.

REEVES, B., LANG, A., KIM, E. Y., AND TATAR, D.   1999.   The effects of screen size and message content on attention and arousal. *Media Psych. 1*, 1, 49–67.

RUDDLE, R., PAYNE, S., AND JONES, D.   1999.   The effects of maps on navigation and search strategies in very-large-scale virtual environments. *J. Exper. Psych. Applied*, 5, 54–75.

SHEPARD, R. N. AND METZLER, J.   1971.   Mental rotations of three-dimensional objects. *Science 171*, 3972, 701–703.

SLATER, M. AND USOH, M.   1993.   Presence in immersive virtual environments. In *Proceedings of the IEEE Conference—Virtual Reality Annual International Symposium*, 90–96.

STREITZ, N. A., GEIBLER, J., HOLMER, T., KONOMI, S., MÜLLER-TOMFELDE, C., REISCHI, W., REXROTH, P., SEITZ, P., AND STEINMETZ, R.   1999.   i-LAND: An interactive landscape for creativity and innovation. In *Proceedings of the CHI 1999 Conference on Human Factors in Computing Systems*, 120–127.

SUZUKI, K. AND NAKATA, Y.   1988.   Does the size of figures affect the rate of mental rotation? *Perception & Psychophysics 44*, 1, 76–80.

SWAMINATHAN, N. AND SATO, S.   1997.   Interaction design for large displays. *Interactions 4*, 1, 15–24.

TAN, D. S.   2005.   Exploiting the Cognitive and Social Benefits of Physically Large Dislpays (Doctoral Dissertation, Carnegie Mellon University, 2004). *Dissertation Abstracts International 65*, 9, 4681.

TAN, D. S., GERGLE, D., SCUPELLI, P., AND PAUSCH, R.   2003.   With similar visual angles, larger displays improve performance on spatial tasks. In *Proceedings of the CHI 2003 Conference on Human Factors in Computing Systems*, 217–224.

TAN, D. S., GERGLE, D., SCUPELLI, P., AND PAUSCH, R.   2004.   Physically large displays improve path integration in 3D virtual navigation tasks. In *Proceedings of the CHI 2004 Conference on Human Factors in Computing Systems*, 439–446.

TAN, D. S., STEFANUCCI, J. K., PROFFITT, D. R., AND PAUSCH, R.   2001.   The Infocockpit: Providing location and place to aid human memory. *Workshop on Perceptive User Interfaces 2001*.

TANI, M., MASATO, H., KIMIYA, Y., KOICHIRO, T., AND FUTAKAWA, M.   1994.   Courtyard: Integrating shared overview on a large screen and per-user detail on individual screens. In *Proceedings of the CHI 1994 Conference on Human Factors in Computing Systems*, 44–50.

THORNDYKE, P. AND HAYES-ROTH, B.   1982.   Differences in spatial knowledge acquired from maps and navigation. *Cog. Psych. 14*, 560–589.

THURSTONE, L. L. AND THURSTONE, T. G. 1941. Factorial studies of intelligence. *Psychometric Monograph*, 2.

TLAUKA, M. 2002. Switching imagined viewpoints: The effects of viewing angle and layout size. *Brit. J. Psych. 93*, 193–201.

WALLER, D., HUNT, E., AND KNAPP, D. 1998. The transfer of spatial knowledge in virtual environment training. *Presence: Teleoperators and Virtual Environments 7*, 2, 129–143.

WRAGA, M., CREEM, S. H., AND PROFFITT, D. R. 2000. Updating displays after imagined object and viewer rotations. *J. Exper. Psych. Learning, Memory, and Cognition*, *26*, 1, 151–168.