Training in Virtual and Real Environments

ROBERT V. KENYON and MICHELLE B. AFENYA

Department of Electrical Engineering and Computer Science, University of Illinois at Chicago, Chicago, IL

Abstract—Transfer of training between real and virtual environments was examined using a pick-and-place task with two different difficulty levels. The task was to minimize the time to move cans from one color coded location in the first row to the same color coded location in the back row and then to reverse the process. In the first task, the front and back disk colors were aligned, and in the second disk order, the front and back disk colors were randomly placed on the table. Subjects trained in one environment were then tested in the other and their performance compared with that of subjects being trained in that environment. Some virtual world-trained subjects showed small but significant improvement in performance compared with the untrained subjects for the real world task for both disk arrangements. The differences in performance between the two groups decreased with trial number until no difference was seen at the end of the sessions. None of the real world-trained subjects showed any significant improvement when performing the task in the virtual world compared with the untrained subjects. These results suggest that transfer-of-training from virtual to real world tasks can take place under certain conditions.

Keywords—Virtual environment, Pick-and-place, Manual control, Transfer-of-Training.

INTRODUCTION

Transfer of training from a synthetic or virtual environment (VE) to the physical world has been an accomplished fact for many years in area of pilot training. Flight simulators (a form of VE) have been used to train commercial and military aviators since World War II and astronauts since the beginning of the space program (10,11). Virtual environments may hold this same potential for more mundane but no less important tasks, such as those in industry.

However, the virtual environment that is used to train pilots has a higher task fidelity than those systems used for other VE applications. This difference in fidelity is attributable to the requirement that, in the most interesting VE applications, one is expected to interact with the environment at close range (*i.e.*, at arms' length). In a simulator, the pilot flies through the environment and the only close range interaction with the visual scene occurs when the plane is close to the ground. Consequently, in most of the VE tasks, interaction with computer generated objects are at a level of detail and scale not generally found in flight simulation. The requirement for close contact with, or direct manipulation of, the virtual surroundings introduces significant problems for training, due to the sensory deprivation subjects experience in VE associated with sound, touch, visual resolutions and interactivity (or update rate) (1,9,6,15,16). However, the sensory information necessary for adequate training can vary from one task to the next and its importance is not fully understood. Therefore, the extent to which VE training can be transferred to the real world needs to be examined.

Recently, Kozak *et al.* (7) examined the transfer of training of a simple pick-and-place task in the virtual world to the real world. Their finding of no transfer from virtual world to real world may show the limits of the head mounted display (HMD) VE system they used for training. However, VE system characteristics can differ significantly depending on the type of system used and this can impact what is learned in the VE. In a HMD system, all visual objects are synthetic. Therefore, subjects cannot see their own hands and body to help provide orientation cues. In our projection-based virtual environment (3), the subject can see the floor and other real objects in the room simultaneously with the synthetic objects making the spatial sense of the virtual objects with respect to the real world, closer to that found in the real world. These differences lead us to hypothesize that in our system, transfer-of-training would occur more readily since the movements of the subject are expected to be more like those in the real world. We replicated their simple pick-and-place task to verify this hypothesis. We also added a task in which the disk arrangement was random rather than ordered. This task was expected to be more difficult to perform and would require the subject to learn a strategy to
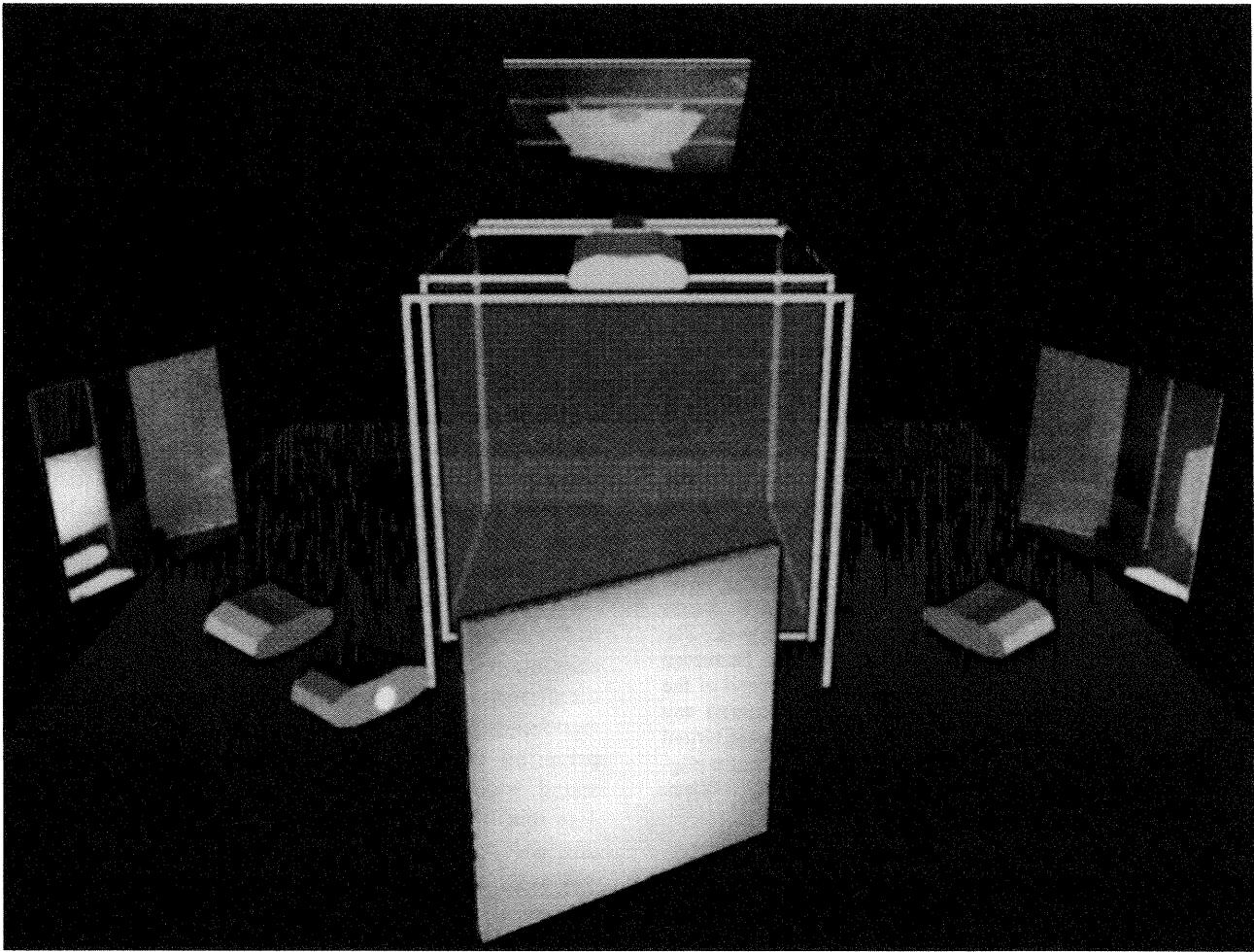
FIGURE 1. A rendering of the CAVE Automatic Virtual Environment (CAVE) used as the virtual world testing and training device.

maximize performance. Consequently, we expected this task to benefit more from training in the virtual world and show more transfer than the simple ordered disk arrangement.

METHODS

Apparatus

The CAVE^TM Automatic Virtual Environment[1]. The virtual world used in this experiment is referred to as the CAVE Automatic Virtual Environment (CAVE) (3). It is a projection-based virtual environment system that surrounds the viewer with four screens. The screens are arranged in a 10-ft cube made up of three rear-projection screens for[2] walls and a down-projection screen for the floor (Fig. 1). Electrohome Marque 8000 projectors are used with P43 coated green tubes to reduce the persistence of the green

phosphor. Each projectors' optics is folded by mirrors due to room size limitations. The images projected onto the CAVE walls are controlled by an SGI Onyx with three Reality Engine 2s. Each Reality Engine is dedicated to rendering the images for one wall of the CAVE limiting the current configuration to two walls (front and left) and the floor.

The design of the CAVE follows an inside-out paradigm, which means that the viewer is inside looking out as opposed to *vice versa* as in the outside-in paradigms of theaters. The CAVE uses window projection where the projection plane and the center of projection relative to the plane are specified for each eye, thus creating an off-axis perspective projection (12). The correct perspective and stereo projections are based on values returned by the position sensor attached to the Stereographics Crystal Eyes stereo shutter glasses. The screen updates at 96 or 120 Hz with a resolution of $1,025 \times 768$ or $1,280 \times 492$ pixels, respectively. Two off-axis stereo projections are displayed on each wall. To give the illusion of 3D, the

---

[1]CAVE is a trademark of the Board of Trustees of The University of Illinois.

viewer wears stereo shutter glasses that enables a different image to be displayed to each eye by synchronizing the rate of alternating shutter openings to the screen update rate. When generating a stereo image, the screen update rate is effectively cut in half due to the necessity of displaying two images for one 3D image. Thus, with a 96 Hz screen update rate, the total image has a maximum screen update rate of 48 Hz. The field-of-view of the CAVE varies between 90° and 120°, depending upon the distance of the viewer from the projection screens.

The CAVE has a second position sensor which is used for the wand: an input device that allows the viewer to interact with the virtual environment. It has three available buttons (normally open switch) and a joystick. The wand uses RS-232 to connect to a Real Time Devices Inc. ADA2710 Analog I/O board. The ADA2710 board is resident in a 486PC and is used to decode the analog and digital inputs that are then sent to the Onyx at a 9,600 baud rate. The wand was not used for this experiment. Instead, the position sensor was removed from the wand and placed on the backside of a glove to track hand movements and to allow more realistic interaction when performing the pick-and-place task.

Head and hand positions are measured with the Ascension Flock of Birds six degree-of-freedom electromagnetic tracker operating at a 30 Hz sampling frequency for a dual sensor configuration. The transmitter is located above the CAVE in the front and has a valid operating range of 7.5 ft. Consequently, movement data is valid only between the top of the CAVE and 1.5 ft above the floor and 8 ft away from the front wall of the CAVE. There are nonlinearities within this range that are caused by the metallic objects and electromagnetic fields created by other devices resident in and about the CAVE. The nonlinearities have been corrected to within 1.5% by linearizing values returned by the position sensing system (4). This linearization is accomplished by using a correction table containing measured positions in the CAVE and then applying linear interpolation to the points that lie between the measured values.

The Glove. A Cross Training glove with the fingers exposed was used to secure the position sensor to the back of the subject's hand. The palm of the glove is made of rubber with the remainder of the glove made of lycra spandex. This gave the glove a snug fit but was not cumbersome to the subject. The wires from the tracker were secured to the subject's arm with tape to ensure that the wires did not interfere with the subject's movements, particularly during the real world testing.

The Grasper. The device used to signal when to pick up the cans in the virtual world was a hand exerciser instrumented to behave as a normally open switch. To close the switch, the grasper had to be compressed (1.5 lb of force)

simulating the grasping of the can in the real world. This information, decoded by the A/D board, was sent to the Onyx for processing in software. The grasper and its cable weighed approximately 4 oz.

The Real World. The experimental equipment consisted of two rows of 3-in colored disks placed on a table measuring 33.5 inches in width and 28 inches in height. Five Coca Cola soda cans were placed on the table upon the front row of disks. The diameter of each of the cans was 2.75 in. The first row of disks were located 4 in from the edge of the table with the second row of disks located 6 in behind it. Disks within each row were 6 in apart. The disks were secured to the table to ensure they would not move during each trial. In each experiment, the front row of disks was arranged from left to right as: red, orange, green, yellow, and blue. For ordered disk arrangement, the back row was organized identically to the front row. For random disk arrangement, the back row disks were distributed from left to right as follows: blue, red, orange, yellow, and green. A yellow square was placed on the table in front of the center can to represent the Start-End position. The experiment was separated into two sets of training/testing blocks. The first set was performed using the ordered disk arrangement and the second using the random disk arrangement. Sand had been placed inside the cans to give each a weight of 4 oz, which was the approximate weight of the grasper and its wire. The subject was outfitted with the stereo shutter glasses and the glove with a position sensor attached to each to record head and hand movements in the real world.

The Virtual World. Subjects were outfitted with the stereo shutter glasses, the glove, and the grasper. The visual cues in the virtual world were designed to match as much as possible those in the real world. The table had a wooden texture on its surface. The objects and their relative distances were scaled to simulate those in the real world. The five soda cans were textured with a Coca Cola logo (Fig. 3). The diameter of the cans was 3 in, and the height and the width of the table were 35 and 64 inches, respectively. A coffee mug textured with a flower print sat on the far end of the table away from the disks and cans. These textured objects were introduced to create as rich an environment as possible, and to provide the subject with strong 2D cues to depth that helped to improve performance (observations from preliminary experiments). However, these enhancements were not without performance penalties. The system update rate fell from 48 to 24 Hz when rendering this complex scene. Finally, the disk colors in the virtual world matched those of the real world with one exception. The orange disk in the real world was replaced with a magenta disk in the virtual world due to the similarity of the yellow and orange colors produced by the projectors in the CAVE. (Note: this change in disk

color did not impact the subjects, since the patterns of the ordered and random disk arrangements were the same in both worlds.)

Since there was a lack of tactile and force feedback correlated to the contact with the cans, a visual cue was needed to provide feedback regarding the position of the hand with respect to the other objects in the virtual world. A red cube served as a 3D cursor to represent the location of the subject's hand in the virtual world. (Note: the subject is able to see his or her own hand as well as the cursor that represents his or her hand in the virtual world.) A can was picked up by placing the cursor inside the can and squeezing the grasper. Similar to picking up a can in the real world, the grasper had to remain compressed in order to move the can in the virtual world. Two cursor conditions within the virtual world were chosen: attached and detached. The attached cursor condition places the cursor in the palm of the subject's hand. The detached cursor condition offsets the cursor's position relative to the subject's hand by 6 inches in both the $-z$ and $+y$ directions. These two conditions were designed to determine if there was a difference between manipulating objects that were proximal and remote to the real hand. The detached condition might have an advantage over the attached condition since there is less possibility of occlusion by the subject's own limbs.

The cursor also served to help the subject deal with other system deficits such as delay, tracker nonlinearity, and positional offsets between the real hand and virtual objects. During rapid hand motion, the subjects' hand position in virtual world lagged behind the hand in the real world due to a 150 ms delay between sensing hand position and its resulting effect on the environment. Without the cursor, subjects could unknowingly close their hand around the virtual can before the computer received the true hand position, resulting in a failure to grasp the object. In another role, the cursor also acts as a reference point or a depth cue. When the cursor is positioned next to an object in the virtual environment, the subject can determine proximity to the object and make appropriate corrections in movement in order to pick up the object. Also, when the real hand is reaching for an object, it could actually occlude the target object by reaching either past the object or through the object. The addition of the cursor helps to determine when and where to grasp the object (2).

Procedure

Each subject was allowed a maximum of 20 min to get familiar with the CAVE environment. The subject was allowed to move about and interact with the application. The familiarization phase did not involve object manipulation. Applications were chosen primarily to make the subject comfortable with the inside-out paradigm of the CAVE and with viewing 3D stereo images. Those subjects who were very familiar with interacting with the CAVE were allowed to waive the familiarization phase.

The subject was positioned on a bench in front of the table and encouraged to remain in this position until the experiment was complete. The task began at the Start-End position and was completed once the subject's hand was placed back in this position at the end of the trial. Starting from the rightmost can, the can was moved from its colored disk and placed on a disk in the back row having the same color. This was repeated for each of the cans. Once all the cans were positioned in the back row, starting with the leftmost can, each of the cans was moved to the front row placing each of the cans on the matching colored disk. When the rightmost can was repositioned in the front row, the subject's hand was placed back at the Start-End position to signal the end of the trial.

The subject was trained and tested on the two disk arrangements. The subjects were randomly distributed into four groups: (i) virtual world with attached cursor group, received training in the virtual world with the cursor located at the hand position and was tested in the real world, (ii) virtual world with detached cursor group, received training in the virtual world with the cursor offset from the hand position and was tested in the real world, (iii) real world with attached cursor group, received training in the real world and was tested in the virtual world with the cursor located at the hand position, (iv) real world with detached cursor group, received training in the real world and was tested in the virtual world with the cursor offset from the hand position. Each subject was instructed not to pass their hand through any of the virtual objects. They were also instructed that this was a test of speed, so they should perform the task as quickly as possible. The subjects were not to sacrifice speed for the sake of accuracy. The trial was accepted if all the cans at least touched the disks. Valid grasps were defined to be when the can was picked up with the hand encircling the sides of the can. In the virtual world, a valid grasp was defined to be when the hand was positioned as if it were encircling the sides of the can.

The subjects were monitored to determine if a trial needed to be repeated due to the subject knocking down or improperly grasping a can, or if a can was placed on the wrong disk color. Each trial was timed using a digital stop watch. The subject was trained by performing 30 trials of the task in the selected training world. The number of trials was selected based on the previous work by Kozak (7). After training, the subjects were tested in the other environment by performing an additional 30 trials of the task. This sequence was used for both disk arrangements. The average time to complete the entire experiment was 2 hr per subject. This included preparation time, subject familiarization time (20 min), breaks between tasks (10

min), and time to complete a block of 30 trials in the virtual world (30 min) and in the real world (15 min).

Subjects

The subjects were limited to the students, faculty, and staff of the university community. The experiment included 24 subjects randomly divided into the four groups mentioned above. All subjects had a great deal of experience interacting with computers. All the subjects were right-handed and had binocular visual acuity. Subjects' ages ranged between 21 to 45 years. All subjects were in good health and had no uncorrected vision problems. Subjects consisted of an even distribution of those naive to the CAVE and those very familiar with the environment.

Analysis

The mean response times for each group were used for the statistical analysis. For both the ordered and random disk conditions, data from groups trained in the real world for the attached and detached cursor conditions were combined into a single population of untrained subjects since the attached and detached conditions were only valid in the virtual environment. A two-way repeated-measures analysis of variance (ANOVA) was used with trial block as the within subjects factor and training (trained/untrained) as the between subjects factor (SPSS for Windows Version 6). Multiple regression analysis (8) was used to test the significance of the slope and intercepts of the regression coefficients using the model: $Y = a_1 + (a_2 - a_1) M + b_1 X + (b_2 - b_1) XM$, where $M = 1$ for trained and 0 for untrained; $X = \log(\text{trial number})$; a_1 and b_1 = intercept and slope for untrained population; a_2 and b_2 = intercept and slope trained population; and Y = completion time.

RESULTS

Completion times for tasks performed in the virtual environment were significantly longer than for tasks in the real world. The average completion times for the ordered and random disk arrangements for all subjects were 19.28 and 21.21 sec respectively in virtual world, and 7.37 and 10.01 sec respectively in real world. The histogram in Fig. 2 shows a typical distribution of completion times for real and virtual world tasks regardless of the disk arrangement used. Completion times in the real world ranged between 6 and 14 sec overall and in the last two 5 trial blocks the average completion time was less than 7 sec. In the virtual world for the ordered disk arrangement, the completion times ranged between 12 and 45 sec, and barely overlapped those from the real world condition. In the last two 5 trial blocks, completion times averaged 16.8 sec. The completion time histogram for random disk arrangement had a similar distribution as in Fig. 2 except each distri-
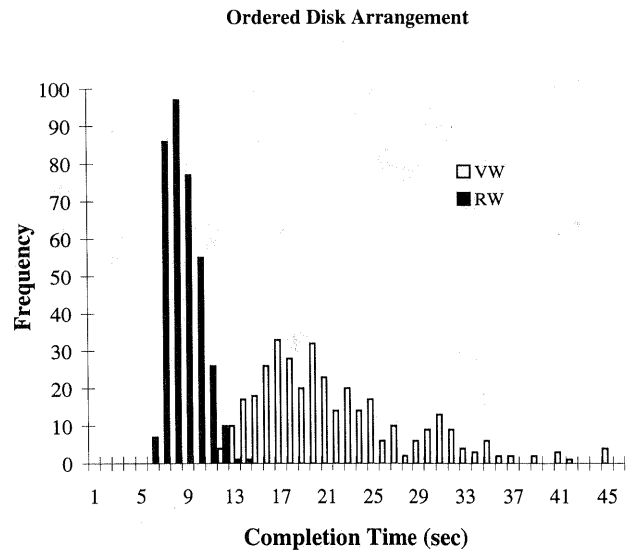
Ordered Disk Arrangement

FIGURE 2. Completion times histogram for performing the ordered disk arrangement task in the real world (RW) and virtual world (VW).

bution was shifted to the right by about 2 sec. The completion times for the random disk arrangement in the real world ranged from 7 to 19 sec with an average time of 9.5 sec for the last two blocks of training. Virtual world completion times were longer with a range between 14 and 50 sec with an average time of 20 sec for the last two blocks of training.

Figure 3 shows overhead and rear views of subject ER's 3D head (red points) and hand motions (white points) while performing the pick-and-place task in the real (upper panels) and virtual worlds (lower panels) using the ordered disk arrangement. To prevent clutter from overlapping data points, only data from the front-to-back portion of task is displayed in the figure. Qualitative differences in head motion amplitude and trajectory are prominently displayed in this figure. Head trajectory in the real world (upper panels) consisted of a smooth curved motion that starts in the center of the frame and then moves forward and to the right to allow the subject to grab the first can. The subject then makes a smooth linear motion from right to left as the cans were moved from near to far. Figure 3A (upper) shows that the subject made very little fore-aft motion. In the virtual world (lower), the subject's head trajectory was larger than that in the real world with a noticeable fore-aft component as the subject moved the cans from near to far. The corresponding hand data (offset from the position of the cans and disks for easier viewing) reflect the large differences in completion times seen in Fig. 2. In the real world condition (Fig. 3, upper), a continuous and rapid flow of hand movement (indicated by the sparse number of data points) from the front to the back is shown. Little hesitation and small vertical displacement in hand movement can be seen in
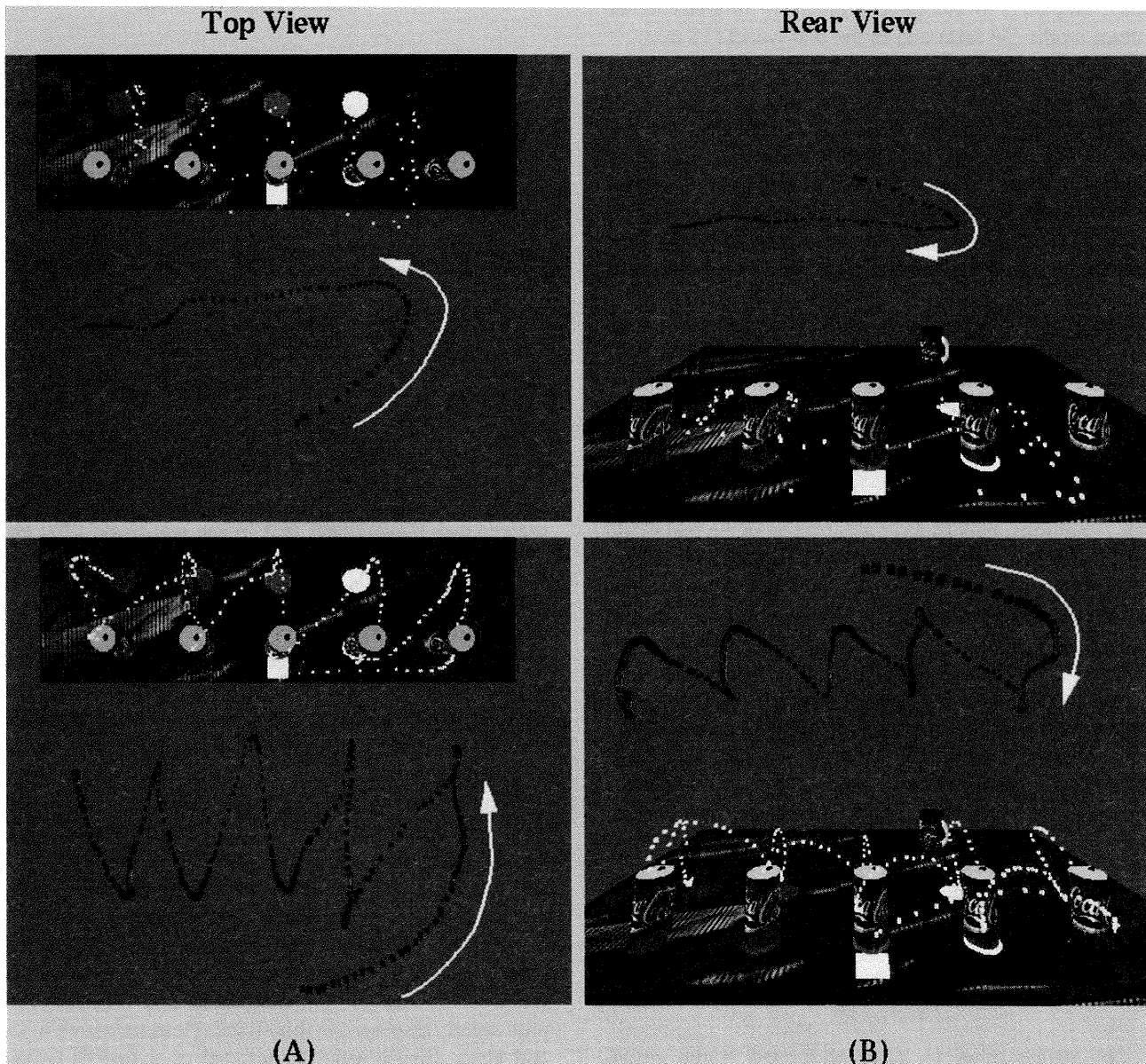
FIGURE 3. Task environment with superimposed head and hand position for the ordered condition in the real world (Top) and the ordered/attached condition in the virtual world (Bottom). To prevent clutter from overlapping data points, only data from the front-to-back portion of task is displayed in this figure. Also, hand data points are offset from the position of the cans and disks for easier viewing. (A) Overhead perspective view of the environment and the motion data to better illustrate the fore-aft motions of the head and hand. (B) Rear perspective view to show both head and hand fore/aft motions, and also the elevation of the hand motion during the task.

these data. In stark contrast, the hand trajectories in the virtual world (Fig. 3, lower), show a high concentration of data points along the path and large vertical displacement in the movements.

The virtual world task took more than twice as long to complete as the real world task. To understand what the increase in completion time may be attributed to, hand position and velocity in the horizontal plane (fore-aft movements) for the ordered/attached condition are plotted

in Fig. 4. The peak velocities in the real world condition (Fig. 4A) are twice as fast as those in the virtual world (Fig. 4B). In addition, the time between zero crossings for movement in the real world (300–400 ms) is less than half that measured in the virtual world (600–1000 ms). The higher concentration of the points at the beginning and end of the can placement indicates longer dwell times when grabbing and releasing the cans in the virtual world.

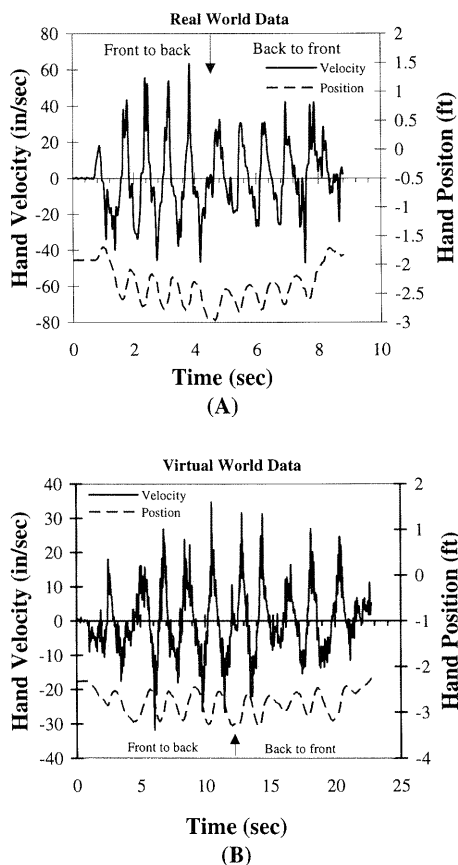Figure 5A shows the mean completion times over

FIGURE 4. Hand velocity and position for subject JM (28th trial) in (A) real and (B) virtual worlds using ordered/attached condition. Front-to-back movement of the cans are indicated by the dividing line at the top or bottom of the graph. Note that the time scale is 2.5 times longer and the velocity scale is half as large in (Panel A versus Panel B). The solid line represents hand velocity and the dashed line represents position.

blocks of five trials. The untrained subjects showed a typical learning curve where the completion time decreased with increasing blocks of trials. For the ordered/attached condition, the mean completion times were all lower than those of the untrained subjects within each block. Despite the lower completion times, a clear improvement in task performance is evident for the trained subjects with increasing block trials. The ANOVA performed on the trained (attached) and untrained subjects' data showed a significant main effect for training ($F[1,8] = 22.05$, $p < 0.002$) and trial blocks ($F[5,40] = 64.96$, $p < 0.001$), but no significant training $\times$ trial block interaction effect was found ($F[5,40] = 0.70$, $p < 0.62$). Consequently, no individual *post hoc* comparisons were performed. For the ordered/detached condition, trained subjects showed lower average completion times compared to the untrained subjects. However, the ANOVA performed on the trained (detached) and untrained subjects' data showed that there was no significant main effect of training ($F[1,8] = 2.88$, $p < 0.128$), but a signif-

icant main effect by trial blocks ($F[5,40] = 80.8$, $p < 0.0001$) and a significant training $\times$ trial block ($F[5,40] = 2.53$, $p < 0.044$) interaction effect. Post hoc t tests revealed that trained subjects' completion times were significantly less than untrained subjects for blocks 2 ($p < 0.021$) and 3 ($p < 0.034$).

Average completion times for all trials for the untrained and the trained (attached) populations are plotted against the log of the trial number in Fig. 5B. These data show that the first trial for the trained and untrained subjects was almost the same, but essentially all other trial times were lower for the trained group for each trial. The regression lines for each group are almost parallel to each other with slopes equal to -0.84 ($R^2 = 0.93$) and -0.72 ($R^2 = 0.985$) for trained and untrained subjects, respectively, but the intercepts differ by 0.7 sec. The t test from the multiregression analysis on these regression line coefficients of slope and intercept revealed significance only for the intercept ($p < 0.001$).

All groups from the random disk arrangement, Fig. 5C, show longer completion times compared with those from the ordered disk arrangement, Fig. 5A, even at the maximum level of training. For the random disk arrangement, both untrained and trained subjects showed improvement in completion times with increasing trial block. The ANOVA on data from the trained (attached) population revealed a significant interaction effect of training $\times$ trial block ($F[5,40] = 3.53$, $p < 0.01$) and a significant main effect of trial block ($F[5,40] = 56.13$, $p < 0.0001$), but no significant main effect of training ($F[1,8] = 1.14$, $p < 0.318$). *Post hoc* t test revealed significance between trained and untrained performance only in block 3 ($p < 0.021$). The ANOVA performed on data from the trained (detached) population showed significance for the main effect of training ($F[1,8] = 5.86$, $p < 0.042$), trial block ($F[5,40] = 48.89$, $p < 0.0001$) and interaction effect of training $\times$ trial block ($F[5,40] = 4.17$, $p < 0.004$). *Post hoc* t test revealed significantly better performance for trained subjects in blocks 3 ($p < 0.003$) only.

The average completion times (Fig. 5D) showed that, although the virtual world trained subjects out-performed the real world subjects by about 1 sec initially, the trial completion times between groups become negligible as the subjects progress to the final trial. The regression lines for each group have slopes of -0.55 ($R^2 = 0.862$) and -0.83 ($R^2 = 0.951$) for trained and untrained respectively, and intercepts that differ by 1.2 sec. The t test from the multiregression analysis on these regression line coefficients of slope and intercept revealed significance for both the slope ($p < 0.001$) and intercept ($p < 0.001$).

Figure 6 shows the mean completion times for the ordered disk condition using the attached cursor in the virtual world for a population initially trained in the real
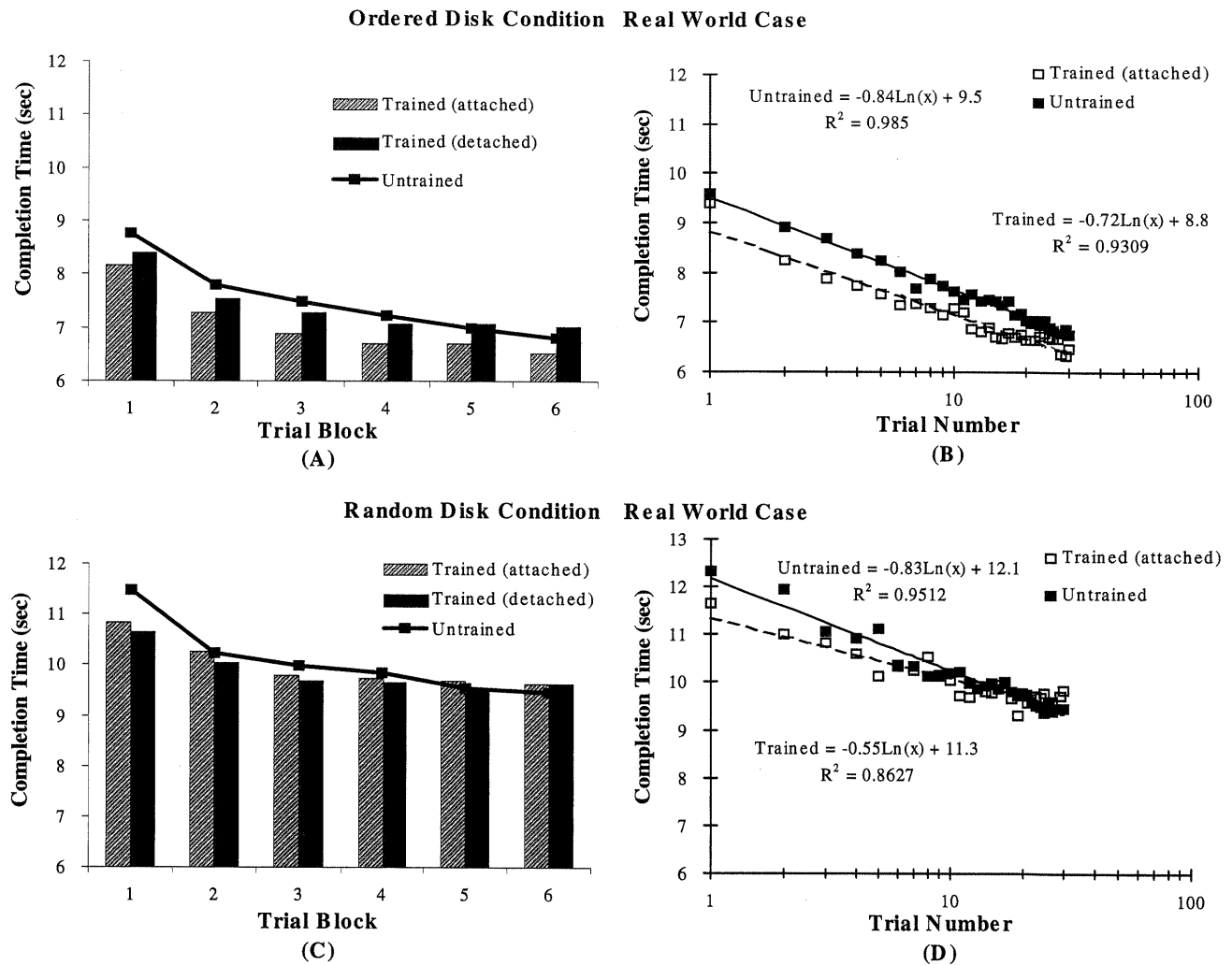
Ordered Disk Condition Real World Case

(A)

(B)

Random Disk Condition Real World Case

(C)

(D)

FIGURE 5. Population mean completion times for experiments performed in the real world. (A) Mean Completion time for trained and untrained groups over blocks of five trials for the ordered disk condition. Bars refer to trained groups and the line to the untrained group. (B) Mean completion times *versus* the log of the trial number for trained (ordered/attached) and untrained populations. The untrained population shown by the solid line and filled symbols and the trained group by the dashed line and open symbols. (C) Mean Completion time for trained and untrained groups over blocks of five trials for the random disk condition. Legend is the same as in Panel A. (D) Mean completion times *versus* log trial number for trained (random/attached) and untrained condition. Legend is the same as in Panel B.

world along with an untrained population. Figure 6A shows the mean completion times over blocks of five trials. The subjects trained in the real world show shorter completion times in the first block of trials than the untrained subjects. However, the ANOVA showed that there was no significant main effect of training ($F[1,8] = 0.04$, $p < 0.839$) nor any interaction effects of training $\times$ trial block ($F[5,40] = 1.34$, $p < 0.268$) but, a significant effect by trial blocks ($F[5,40] = 36.65$, $p < 0.0001$). None of the other virtual world cases showed any significance for training or interactive effects. Significant main effect of trial block was found for all conditions.

Average completion times for all trials for the untrained and the trained (detached) populations are plotted against the log of the trial number in Figure 6B. The average

completion times show a 5 sec difference between the trained and untrained groups in the first two trials, but the untrained group rapidly reduced this difference with subsequent trials. The regression lines for each group have slopes of -3.68 and -5.24 for the trained and untrained populations, respectively, which are much larger than those found for the real world cases in Fig. 5. Both the slope and intercepts between the virtual world trained and untrained populations were found to be significantly different from each other ($p < 0.0001$).

DISCUSSION

These experiments have shown that a task learned in a virtual world can improve performance on the same task

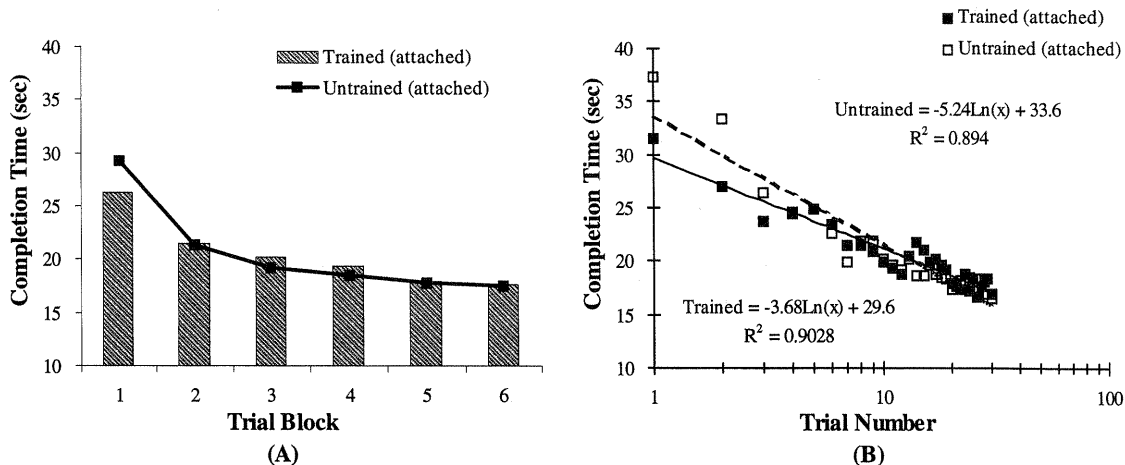**Ordered Disk Condition   Virtual World Case**

FIGURE 6. Population mean completion times for experiments performed in the virtual world. (A) Mean Completion time for trained and untrained groups over blocks of five trials for the ordered/attached condition. Legend is the same as in 5(A). (B) Mean completion times *versus* log trial number for trained and untrained order/attached population. Legend is the same as in Fig. 5B.

when it is performed in the real world. However, the transfer-of-training from the virtual to the real world is not 100%, and in most cases the significance appeared briefly during the experiment. These results differ from those reported by Kozak, who found no significant performance change from the virtual to the real world. Kozak's subjects' completion times averaged 63 sec for the virtual world and 5.9 sec in the real world compared with 19.2 and 6.2 sec, respectively, in these experiments using the same disk arrangement. One might expect that, as the completion times moved closer to those in real world, the transfer of skills learned in the virtual world to real world would improve.

Comparing this study to Kozak's is very difficult due to the significant differences in experimental conditions and equipment. An obvious difference between these two studies is in virtual world characteristics. In Kozak's HMD system, subjects are unable to view simultaneously, both real and virtual objects in the environment. Therefore, they cannot see the spatial relationship between their real hands, arms, and body with respect to the virtual world. Furthermore, if the mapping of the synthetic hand in the virtual world is not sufficiently aligned with the hand in the real world, the subject must adapt to the perceived location of the hand in the virtual world with respect to the actual kinesthetic/proprioception of the limb (5,14). Welch (13) has shown that adaptation to prismatic displacement of the hand in a manual pointing task can take as many as 35 trials to produce 86% adaptation. This adaptation time may have a significant effect on performance and learning of spatially dependent behavior that is needed to perform the task when training time is limited.

The large difference in completion times between our real world and virtual world data may reflect a true dif-

ference in motor control movement between the two environments. Our own observations of subjects' behaviors in each environment were that subjects' movements were more deliberate in the virtual world than in the real world (Fig. 3). The similar shape but lower peak velocities for virtual world task movements compared with those in the real world (Fig. 4) may be indicative of the deliberateness of the movements in the virtual world. These changes in movement dynamics in the virtual world may have resulted from the poverty of the sensory feedback available in that environment. The virtual world does not contain shadows, object weight, tactile cues, occlusion of the hand with the object, and normal temporal/spatial conditions. Two of the more important considerations from this list are the lack of a tactile cue when grabbing the can and the 150-ms delay in the image generating system. Subject interviews revealed that, in the real world, vision was supplemented by the tactile cues when moving the cans. In the virtual world, subjects must always use visual feedback to confirm contact with the can and so the actions must be more deliberate than in the real world. The delay between true hand movement and that of the cursor within the virtual world also contributes to the deliberateness of the movements. Subjects learned on the first trial that the position of the cursor lagged behind that of their hand during high velocity movements, and so they adjusted their control to match the dynamics of the system. We believe that these two factors are the main contributors leading to the differences found in motor control in the two environments.

We initially believed that a task requiring the subject to learn a strategy to perform the task well would show a high percentage of transfer from one world to another. The ordered disk arrangement required more motor learning

than strategy to improve performance. In the random disk arrangement, we hypothesized that the strategy component would out weigh the motor component and therefore show a high transfer. The average completion time data showed that the random arrangement took more time to complete in both the virtual and real worlds than the ordered arrangement, indicating that it was more difficult to perform. Also, the data in Fig. 5C and D showed that there were lower completion times for the trained population for the first few trials, but only the third block of training showed any significance between trained and untrained subjects. We speculate that the advantage of learning a strategy for moving cans for the random disk pattern in the virtual world was less important than learning the proper motor coordination to perform the task. For example, large crisscross movements of the hand were needed to move the cans to their assigned locations. Perhaps, learning these more complex motor patterns is more difficult given the sparse sensory cues available in the virtual world.

Our finding of some transfer-of-training from the virtual to the real world is encouraging for the use of virtual environments to improve real world performance. However, this result is not very robust in that it was not seen across all conditions tested. Given the difficulties in producing virtual world sensory parity with the real world, any large transfer-of-training from one to the other may not be found until the fidelity for task duplication between the two worlds is better. For example, tasks that are mainly tests of speed may not transfer well due to the current sensory mismatch between the virtual and real world. Further studies will be needed to better understand the very complex relationship of what characteristics are important for particular kinds of tasks.

In the flight simulation area, pilots spend many hours in the simulator learning to fly. However, simulators have a higher sensory fidelity for flight training than current virtual world systems have for arms' length interactions with synthetic objects. Our subjects were trained in the virtual world for only a half an hour before being tested in the real world. Increased training time in the virtual world may show a large positive transfer to the real world task if the training time is selected appropriately for the task being performed. However, the learning curve data from subjects operating in the virtual world show a clear asymptotic curve that approaches 17 sec (Fig. 6). It seems unlikely, for this task at least, that more training time would result in any further significant performance improvement in the completion times in the virtual world. Since these data remain distant from the completion times for the real world task, we speculate that only small improvements in transfer would be expected with further training. Furthermore, if the characteristics of the virtual world greatly mismatch those needed to successfully perform the task in

the real world, additional training could result in a negative transfer-of-training; a problem that seriously concerns the flight simulation community.

We found no significant transfer or lasting performance improvement from the real world to the virtual world, indicating that any gain in task training is masked by the need to adapt to the virtual world's altered sensory cues. In comparing the learning curves between untrained populations in the real and virtual worlds, the virtual world provided a greater learning experience for our subjects than did the real world, as attested by the huge difference in the slopes between the two. The virtual world task showed a change of 10 sec between the first and last trials, whereas the real world task showed only 2 sec difference. Clearly, the large differences in the sensory environment between the two conditions appear to be responsible for these large gains in learning in the virtual world.

No significant changes in performance could be found between subjects trained using an attached or detached cursor. However, the question of whether interaction with the virtual world may best be performed with virtual or real limb is an interesting, unresolved question. In environments where both real and virtual objects are viewed simultaneously, the trade-off between true hand movement and related hand movement has to be addressed. For example, when subjects view a virtual scene the optical power of their eye (accommodation) is set for a certain clear vision distance related to the power of the light rays from the fixed optics of the image display system (such as the distance to the projection screen). When a real object, such as one's hand, enters the scene, its accommodative stimulus may be very different from that of the neighboring virtual object(s) that are in focus. The difference in accommodative stimulus between the real and the virtual objects now poses a forced choice condition on the subject: make the virtual object clear and therefore the real object blurred or *vice versa*. The advantages of seeing ones own hand in the environment must be weighed against this and other issues for each task that is used in VE. The task used in these experiments allowed the subject little time to view both hand and virtual object since the task was one of speed of motion. In other tasks, in which such interactions are more frequent and for longer periods of time, there may be a significant difference between these two modes of virtual world interaction.

One of the conclusions from this research is that current VE technology should be applied carefully to training situations if it is to have positive impact on training motor activities. The results we obtained here, although small, showed some significant performance improvement with training. However, the impact of poor sensory cues on performance in VE is also evident in our data. Addition of audio and tactile cues to the environment would greatly

improve the VE experience and reduce the subject's strict reliance on visual cues. Our subjects were also faced with the need to adapt to the large system delays inherent in our system. Reducing system latencies in the tracking system and intersystem process communication would add to the realism of the experience. Finally, a richer virtual environment might include shadows, more realistic textures, and changes in the virtual object's optical stimulus to accommodation. The impact of each of these specific visual scene features on a subjects performance in VE is currently under investigation in our lab.

Finally, training in a virtual environment is not free, even when the environment already exists. The value of training in a VE might be determined by assessing the amount of time needed to train; its associated costs in personnel, maintenance, and equipment; the safety benefits for VE training; and the amount of positive transfer that takes place between the two environments. Clearly, to justify the costs choosing a task that would give the greatest transfer from the virtual to the real world is important. Also, the more people that can be trained per session, the more cost-effective will be the use of a VE. Although not investigated here, for some tasks, passive viewing rather than active participation in the task might be an effective means to transfer strategic skill from the virtual to real world (similar to athletes watching game films to prepare for next week's opponent). In such cases, many people can be trained simultaneously using systems like the CAVE where more than one person can stereoscopically view the scene simultaneously. Additional research on how subjects learn in virtual environments for various types of tasks would help us better understand how people adapt to these new environments.

REFERENCES

1. Bishop, G., W., Bricken, F. Brooks, M. Brown, C. Burbeck, N. Durlach, S. Ellis, H. Fuchs, M. Green, J. Lackner, M. McNeill, M. Moshell, R. Pausch, W. Robinett, M. Srinivasan, I. Sutherland, R. Urban, and E. Wenzel. Research directions in virtual environments. *Computer Graphics* 26N3:153–177, 1992.
2. Bos, P. J. Performance limits of stereoscopic viewing systems using active and passive glasses. Proc. of the IEEE Annual Virtual Reality Int. Symposium (VRAIS), Seattle, WA, Sept. 18–21, 1993, pp. 371–376.
3. Cruz-Neira, C., D. J. Sandin, and T. A. Defanti. Surround-screen projection based virtual reality: the design and implementation of the CAVE. *Computer Graphics* 27:135–142, 1993.
4. Ghazisaedy, M., D. Adamczyk, D. J. Sandin, R. V. Kenyon, and T. A. Defanti. Ultrasonic calibration of a magnetic tracker in a virtual reality space. Proc. of the IEEE Annual Virtual Reality Int. Symposium (VRAIS), Raleigh, NC, Sept. 1995, pp.179–188.
5. Held, R., and A. Hein. Adaptation to disarranged hand-eye coordination contingent upon re-afferent stimulation. *Percept. Motor Skills* 8:87–90, 1958.
6. Ishii, M., and M. Sato. A 3D interface device with force feedback: a virtual work space for pick-and-place tasks. Proc. of the IEEE Annual Virtual Reality Int. Symposium, Seattle, WA, Sept. 18–21, 1993, pp. 331–335.
7. Kozak, J. J., P. A. Hancock, E. J. Arthur, and S. T. Chrysler. Transfer of training from virtual reality. *Ergon* 36:777–784, 1993.
8. Montgomory, D. C., and E. A. Peck. *Introduction to Linear Regression Analysis*. New York: Wiley and Sons, 1982, pp. 109–180.
9. Ohzu, H. Artificial 3-D displays and visual functions. *Vision Sci. Appl.* (Conf. ed.) 2:48–51, 1994.
10. Rolfe, J. M., and K. J. Staples. Flight simulation. Cambridge, UK: Cambridge University Press, 1986, pp. 232–249.
11. Schachter, B. J. Computer image generation. New York: Wiley-Interscience, 1983, pp. 187–220.
12. Schmandt, C. Spatial input-display correspondence in a stereoscopic computer graphics workstation. *Computer Graphics* 17N3:253–259, 1983.
13. Welch, R. Prism adaptation: The "target-pointing effect" as a function of exposure trials. Percept. Psychophysiol. 9:102–104, 1971.
14. Welch, R. Adaptation of space perception. In: *Handbook of Perception and Human Performance,* vol 1., edited by K. Boff, L. Kaufman, and J. Thomas. New York: Wiley-Interscience, 1986, pp. 24-1–24-37.
15. Wickens, C. D., and P. Baker. Cognitive issues in virtual reality. In: *Virtual Environments and Advanced Interface Design,* edited by W. Barfield and T. Furness. New York: Oxford University Press, 1995, pp. 514–541.
16. Zhai, S., and P. Milgram. Human performance evaluation of manipulation schemes in virtual environments. Proc. of the IEEE Annual Virtual Reality Int. Symposium (VRAIS), Seattle, WA, Sept. 18–21, 1993, pp. 155–161.