Albert A. Rizzo
arizzo@usc.edu
Integrated Media Systems Center
University of Southern California
Los Angeles CA 90089-2561

J. Galen Buckwalter
Galen.X.Buckwalter@kp.org
Southern California Permanente
Medical Group
Dept. of Research and Evaluation
Pasadena, CA 91101

Jocelyn S. McGee
Todd Bowerly
tbowerly@fuller.edu
Fuller Graduate School of
Psychology
Pasadena, CA 91101

Cheryl van der Zaag
vanderz@usc.edu
School of Gerontology
University of Southern California
Los Angeles CA 90089-0191

Ulrich Neumann
Integrated Media Systems Center

Marcus Thiebaux
thiebaux@isi.edu
Information Sciences Institute
University of Southern California
Marina del Ray, CA 90292

Laehyun Kim
Jarrell Pair
Clint Chua
Integrated Media Systems Center

Virtual Environments for Assessing and Rehabilitating Cognitive/Functional Performance

A Review of Projects at the USC Integrated Media Systems Center

Abstract

Virtual reality (VR) technology offers new options for the creation of sophisticated tools that could be applied in the areas of assessment and rehabilitation of cognitive and functional processes. VR systems allow for the precise presentation and control of dynamic, multisensory, three-dimensional (3-D) stimulus environments, as well as the recording of all behavioral responses that occur within them. Assessment and rehabilitation scenarios that would be difficult if not impossible to deliver using conventional neuropsychological methods are now being developed that take advantage of these VR assets. If empirical studies demonstrate effectiveness, virtual environments (VEs) could be of considerable value for better understanding, measuring, and treating persons with impairments due to traumatic brain injury, neurological disorders, and learning disabilities. This article describes the progress of a VR research program at the USC Integrated Media Systems Center and Information Sciences Institute that has developed and investigated the use of a series of VEs designed to target (i) molecular visuospatial skills using a 3-D, projection-based ImmersaDesk system, and (ii) attention (and soon memory and executive functioning) processes within ecologically valid functional scenarios utilizing a head-mounted display (HMD). Results from completed research, rationales and methodology of works in progress, and our plan for future work is presented. Our primary vision has been to develop VR systems that target cognitive processes and functional skills that are of relevance to a wide range of patient populations with central nervous system (CNS) dysfunction, as well as for the assessment of unimpaired performance. We have also sought to select cognitive/functional targets that intuitively appear well matched to the specific assets available with currently available VR technology.

1 Introduction

Virtual reality (VR) is rapidly evolving into a pragmatically usable technology. As continuing advances unfold in underlying enabling technologies in the areas of computing power, graphics and image capture, display systems, interfacing tools, immersive audio, haptics, wireless tracking, voice recognition, intelligent agents, and VR authoring software, more useful and usable virtual environments (VEs) are becoming possible. Concurrent with these

Presence, *Vol. 10, No. 4, August 2001*, 359–374

technological advances, clinicians and researchers are beginning to recognize VR's potential as a new tool for the study, assessment, and rehabilitation of cognitive processes and functional abilities (Brown, Kerr & Wilson, 1997; Pugnetti et al., 1995; Rizzo & Buckwalter, 1997; Rose, 1996). Much like an aircraft simulator serves to test and train piloting ability, virtual environments (VEs) can be developed to present simulations that can be used to assess and rehabilitate human cognition and behavior. The capacity of VR technology to create dynamic, multisensory, three-dimensional (3-D) stimulus environments within which all behavioral responding can be recorded offers clinical tools that are not available using traditional neuropsychological methods. Individuals who may benefit from these applications include persons with cognitive and functional impairments due to traumatic brain injury (TBI), neurological disorders, and developmental/learning disabilities. In this regard, a growing number of laboratories are developing research programs investigating the use of VEs for these purposes, and initial exploratory studies reporting encouraging results have begun to emerge (Rizzo, Buckwalter, & van der Zaag, 2001). This work has the potential to advance the scientific study of normal cognitive and behavioral processes, and to improve our capacity to understand, measure, and treat impairments in these areas that are typically found in clinical populations with central nervous system (CNS) dysfunction.

This evolving application area also fits well within the context of the "Information Society for All" concepts that have recently been addressed in the human-computer interaction literature (Stephanidis et al., 1999). Efforts in this area support the development of computer and information technology (CIT) that accommodates the broadest range of human abilities, skills, requirements, and preferences. With the emergence of a global information society, increasingly based on the production and exchange of information, those who are able to thoughtfully develop and apply more useful and usable CIT will be in a position to impact fundamental changes for advancing human welfare. The potential benefits of this paradigm shift for those with special needs could occur as VE systems redefine the assessment and rehabilitative strategies that are used with clinical populations.

Now being developed and tested are VR applications that focus on component cognitive processes including attention, executive functions, memory, and spatial abilities. Functional VE training scenarios have also been designed to test and teach instrumental activities of daily living such as crossing a street, driving an automobile, preparing meals, shopping in a supermarket, using public transportation, and navigating in a wheelchair. More involved discussion of the rationales, issues, and applications of VR for neuropsychological targets can be found in a recent review chapter (Rizzo et al., 2001). These initiatives have formed a foundation of work that provides support for the feasibility and potential value of further development of neuropsychological VE applications. The initial success of these VE scenarios give hope that the twenty-first century will be ushered in with new and useful tools to advance a field that has long been mired in the methods of the past. As well, major funding agencies in the United States have realized the potential value of effort in these areas. For example, on a more global level, the National Science Foundation stated in a recent VR program announcement that "computer simulation has now joined theory and experimentation as a third path to scientific knowledge. Simulation plays an increasingly critical role in all areas of science and engineering . . . " (NSF PA# 98-168, p.1). More specific to the application of VR in neuropsychology, the National Institute on Disability and Rehabilitation Research (NIDRR) in a recent position statement highlighted that "the benefits of combining virtual reality with rehabilitation interventions are potentially extensive" and specifically calls for research "to determine the efficacy of virtual reality techniques in both rehabilitation medicine and in applications that directly affect the lives of persons with disabilities" (NIDRR Web page).

In view of these issues, this article focuses on the progress of a VR research program at the University of Southern California that has developed and investigated the use of a series of VEs that have been designed to target the assessment and rehabilitation of cognitive and functional processes. Results from completed research, rationales and methodology of works in progress, and our plan for future work is presented. Our primary vision has been to develop VR systems that target cognitive processes and functional skills that are of relevance to a wide range of patient populations with CNS dysfunction. We have also sought to select cognitive/functional targets that intuitively appear well matched to the specific assets that exist with currently available VR systems. Consequently, we have evolved two parallel programs: (i) the targeting of molecular visuospatial skills using a 3-D projection-based ImmersaDesk system, and (ii) the targeting of attention (and soon memory and executive functioning) processes within ecologically valid functional scenarios using a head-mounted display (HMD) system. At the time of the initial development of these applications, we used SGI Onyx systems to produce our scenarios. However, we have begun the process of transferring the scenarios to a PC platform in view of the rapid advances that have occurred in this area lately and due to the desire to develop more-economical systems that could reach a wider range of potential users.

2 Visuospatial Applications Using the ImmersaDesk Projection Display System

A component-based approach was used to address visuospatial ability through the use of a suite of ImmersaDesk-delivered 3-D applications. (See figure 1.) Using this display system, a series of VE scenarios have targeted mental rotation (MR), depth perception, field dependence (3-D rod and frame test), static and dynamic manual tracking, and visual field specific reaction time. These scenarios were designed to leverage the 3-D interactive assets available with this type of projection-based system in the development of a series of tasks that could assess and possibly rehabilitate these more molecular components of visuospatial functioning.

Visuospatial ability is an important cognitive domain to consider in the assessment of neurological disorders, traumatic brain injury, and neuropathological conditions of aging. For example, spatial orientation abilities are an important variable in the differential diagnosis of dementia. Research indicates that victims of Alzheimer's disease have an 84% incidence of spatial orientation impairments compared to only a 4% incidence in frontotemporal dementia (Miller et al., 1997). Impairments in spatial orientation were also shown to be more common in Alzheimer's disease compared to both normal elderly and those with vascular dementia (Gainotti, Parlato, Monteleone & Carlomagno 1992; Signorino et al., 1996). Similar impairments have been observed following the occurrence of traumatic brain injury and stroke (Lezak, 1995). Tests of spatial ability, including the MR variable, are also commonly used for the study of brain/behavior relationships, particularly regarding sex differences in cognition. MR ability produces the most consistent and sizeable sex differences, in favor of males, in the cognitive literature (Voyer, Voyer, & Bryden 1995). Consequently, a lively body of work has emerged examining MR, as well as with other cognitive variables in which female advantages appear (such as verbal fluency and fine motor skills, among others). Studies have reported differential cognitive performance due to such hormonal factors as onset of menopause, estrogen and testosterone administration, and stage of the menstrual

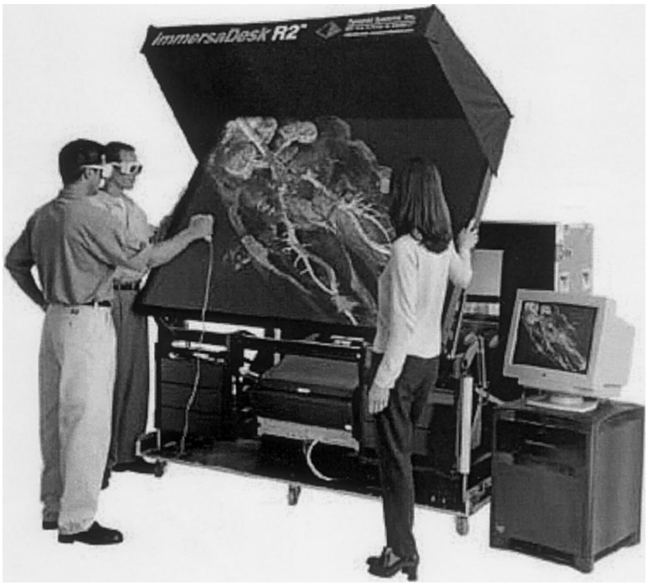

Figure 1. *ImmersaDesk 3-D projection display system.*

cycle (Gouchie & Kimura, 1991; Kampen & Sherwin, 1994; Silverman & Phillips, 1993). However, these findings remain controversial. Several studies have attempted to explain cognitive sex differences as the product of sociocultural influences, and on nonspecific testing performance factors related to the use of timed tests and "reluctance to guess" factors (Richardson, 1994; Qubeck, 1997; Delgado & Prieto, 1996). It has also been suggested that the effect size in gender differences has been decreasing in recent years. However, meta-analytic research has argued against these conclusions (Masters & Sanders, 1993; Voyer et al., 1995). These issues (which are ongoing research interests at our affiliated lab at the USC Alzheimer's Disease Research Center) and our interest in the potential usefulness of VR motivated our development of the Virtual Reality Visuospatial Skills Project.

2.1 Study I: Young Adults

Our initial investigation focused on the development of a VE for the study, assessment, and rehabilitation of the visuospatial ability referred to as *mental rotation* (MR). MR is a well-studied visuospatial variable, which can be described as a dynamic imagery process that involves "turning something over in one's mind" (Shepard & Metzler, 1971). Everyday life situations rely on this ability to use imagery to turn over or manipulate objects mentally. These include automobile driving judgments, organizing items in limited storage space, using a map, sports activities, and many other situations in which one needs to visualize the movement and ultimate location of physical objects in 3-D space. High-level mathematics performance has also been linked, in large part, to MR ability (Casey, Nuttall, Pezaris, & Benbow, 1995). Indeed, in a recent *Los Angeles Times* interview, it was noted that the world-renowned physicist Stephen Hawking "translates mathematics into geometry, and turns around geometrical shapes in his head" (Cole, 1998).

The initial MR investigations began almost thirty years ago with the work of Shepard and Metzler (1971), who tachistoscopically presented pairs of two-dimensional perspective drawings to subjects and required them to make judgments as to whether the 3-D objects they portrayed were the same or different. (See figure 2.) A near perfect linear relationship was found between the amount of angle rotation difference between the pairs of objects, and the reaction time to decide whether or not the objects were the same or different. Since precise mathematical relationships between hypothesized mental representations and behavioral performance are relatively rare, MR has been the focus of considerable research. Traditional 2D measures for the assessment of mental rotation have produced intriguing findings, yet lack the precision needed to better understand this spatial ability. The most common test uses two-dimensional image stimuli that portray three-dimensional objects and requires mental processing of the stimuli without any motoric involvement (Vandenberg & Kuse, 1978).

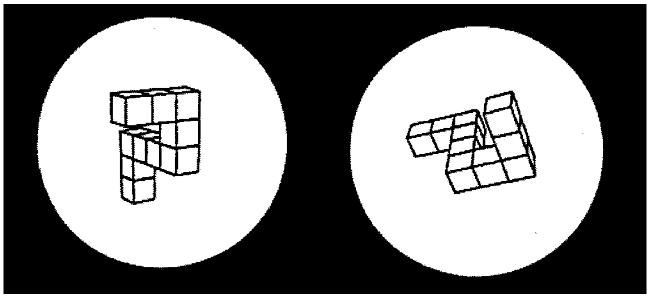

Figure 2. *Mental rotation stimuli.*

In our initial study, MR was targeted via a manual VR spatial-rotation task (VRSR) that required subjects to manipulate and superimpose block configurations within a VE. The use of VR for the assessment of visuospatial abilities was envisioned to allow for greater control and description of 3-D stimuli along with more precise measurement of responses. This was expected to support a more accurate characterization of the cognitive processes that are involved in functional visuospatial skills than is afforded by standard 2-D, paper-and-pencil measures. In addition, the examination of changes in spatial performance following VRSR training was anticipated to provide insight for the development of potential rehabilitation strategies. In the first project with young healthy adults, we investigated the following research questions:

1. Are there negative side effects associated with the use of this VE?
2. Does performance on the spatial-rotation VE demonstrate adequate psychometric properties?
   a. Is the VRSR task reliable in terms of internal consistency (coefficient a) and test-retest measures?
   b. Does the VRSR task demonstrate concurrent validity in the pattern of associations observed with standard visuospatial tests?
3. Do the same sex differences that appear on the pencil-and-paper MRT appear on spatial-rotation VE performance?
4. Does VRSR performance improve with practice (100 training trials) as seen by comparing twenty pre-training VR MRT items with twenty identical post-training VR MRT items (intra-method generalization)?
5. Does VRSR training improve post-training pencil-and-paper MRT performance?

2.1.1 Method

2.1.1.1 Subjects. Sixty subjects (26 males and 34 females) between the ages of eighteen and forty were tested. Subjects included employees recruited at the Information Sciences Institute of the University of Southern California, graduate students from the Fuller Graduate School of Psychology, and undergraduate students from the University of Southern California and California State University at Los Angeles.

2.1.1.2 Virtual Reality System and Procedure. The Virtual Reality Spatial Rotation (VRSR) system uses an ImmersaDesk, drafting-table format, rear-projection display. The Pyramid Systems ImmersaDesk employs stereo glasses and magnetic head and hand tracking. This large-screen display system offers a type of VR that is semi-immersive. It features a 4 ft. by 5 ft. rear-projected screen positioned at a 45 deg. angle. The size and position of the screen give a wide-angle view that affords the ability to look down as well as forward.

The VRSR system is designed to present a target stimulus that consists of a specific configuration of 3-D blocks within a virtual environment (similar to figure 2). The stimuli appear as "hologram-like" three-dimensional objects floating above the projection screen. After presentation of a target stimulus, the participant is presented with the same set of blocks (control object) that needs to be rotated to the orientation of the target and superimposed on it. The participant manipulates the control object by grasping and moving a sphere-shaped "cyberprop," which contains an Ascension Flock of Birds tracker. The motion of the sphere is imparted upon the control object. Upon successful superimposition of the control and target objects, a "correct" feedback tone is presented, and the next trial begins. The new control object appears attached to the sphere (user's hand), and the new target appears a short distance away. In this mode of interaction, users do not need to press any buttons or select any objects. Control objects simply appear attached to the sphere for users to manipulate. We calculate the following information on the stimuli. The orientation of a stimulus can be represented by a single linear rotation around a three-dimensional vector. We specify orientation for each stimulus as a group of four values: three defining a three-dimensional vector of unit length and an angle in degrees. We define magnitude as the angular difference between two orientations. The stimulus orientation can be aligned to the target orientation by a single minimal rotation around some fixed three-dimensional vector. The degrees required to do this is the magnitude of the rotational task, ranging from 0 deg. to 360 deg. If the rotational axis is nearly parallel to the viewer's line of sight, rotations of the object will not reveal new faces to the viewer, and the task is equivalent to a 2-D rotational task. As the axis is moved to become parallel to the viewplane, the task requires a fully 3-D understanding of the object's appearance, and this is considered to be a task of high complexity. This is calculated as the sine of the angle between the rotational axis and the line of sight, and so varies from 0 to 1.

We presently have assessed rotational ability by recording the amount of time to complete the rotation as well as the efficiency of the solution. The most efficient execution of a rotational task follows the shortest angular path from stimulus to target, about a fixed axis. Samples are taken at regular intervals during task execution, defining an angular path as the subject searches for a solution. Summing the angular differences between sequential samples gives the angular length of that path, and calculating the ratio of the shortest possible path to this summed length gives the efficiency of the task execution. This value varies from > 0 (poor) to 1 (ideal).

The experimental sessions took place over a two-hour period. After informed consent was obtained, basic demographic information, computer experience and usage, and spatial activities history (Newcombe, Bondura & Taylor, 1983) was recorded. Next, a baseline measure of mental rotation ability is assessed using a redrawn version (Peters et al., 1995) of the Mental Rotation Test (MRT-A) of Vandenberg & Kuse (1978), a twenty-item, two-dimensional, paper-and-pencil task. Subjects then complete a comprehensive neuropsychological battery administered under standard conditions. The neuropsychological testing battery included a diverse collection of instruments in addition to the MRT-A. Verbal attention and mental control was assessed with the Digit Span Forward and Backward test from the Wechsler Adult Intelligence Scale-Revised (Wechsler, 1981). Visuoconstruction abilities were measured by the Block Design subtest of the WAIS-R. The Trail-Making Tests A and B was used to evaluate executive control processes and attention (Army Individual Test Battery, 1944). The Judgment of Line Orientation test was used to evaluate visuoperceptual skills (Benton, Varney, & Hansher, 1978). The California Verbal Learning Test was employed to assess verbal learning and memory (Delis, Kramer, Kapler, & Oper, 1983). Visual (nonverbal) memory was evaluated by the Visual Reproduction subtest of the Wechsler Memory Scale-Revised (Wechsler, 1987). These tests are all commonly used for neuropsychological assessment of these cognitive processes and, as such, have considerable normative data available.

Following the completion of the neuropsychological battery, subjects completed the Motion History Questionnaire (Kennedy & McCauley, 1984) and Simulator Sickness Questionnaire (Kennedy, Lane, Berbaum, & Lilienthal, 1993), which included a pre-VE exposure symptom checklist. Experimental subjects then participated in the fifteen-minute VRSR task that allowed for both testing and training of spatial rotation abilities. After five nonrotational practice trials, each subject's VRSR baseline performance was assessed over twenty trials using a VE version of the items from the pencil-and-paper MRT. Next, 100 training trials of increasing stimulus complexity were administered. After a one-minute break, the original twenty VRSR trials were administered again to measure changes in performance following VE training. VE data was quantified by analyzing time to completion per trial, path efficiency, and various compiled measures, such as the total time for the first twenty VRSR items versus the last twenty. Control subjects were given a filler task (crossword puzzle) of matching duration instead of the VE exposure. The Simulator Sickness Questionnaire, which contains a post-VE exposure symptom-checklist, was then given to each subject. Finally, an alternate form of the paper-and-pencil MRT was administered to assess changes in mental rotation performance.

2.1.2 Results. A number of encouraging findings emerged, including minimal self-reported VE-related side effects, reasonable psychometric properties of the VRSR test, provocative relationships with standard NA tests, a lack of gender differences compared to the pencil-and-paper performance, training improvement, and significant transfer of training with low initial MRT pencil-and-paper performers.

2.1.2.1 Side Effects. An analysis of individual symptoms from the SSQ symptom checklist found that, for the VRSR group, only "blurred vision" showed a significant increase from pre- to post-testing ($F(1,54) = 3.99$; $p < 0.05$). No significant increase in the overall SSQ symptom checklist total was found for the VRSR group.

2.1.2.2 Reliability. The VRSR reliability was in the moderate range on calculations of internal consistency and matched parcel reliability. Typically, neuropsychology instruments boast reliability coefficients ranging from 0.80 to 0.95 (Mitrushina, Boone, & D'Elia, 1999). Sattler (1988) asserts that a neuropsychology test's reliability coefficient must approximate or exceed 0.80 in magnitude, and that coefficients of 0.90 or above are considered the most desirable. The VR MRT's reliability fell below this range with a coefficient γ equal to 0.71 on its Speed Index and 0.73 for the Efficiency Index. Cronbach's alpha for these indices were 0.65 and 0.69, respectively. However, our reliability findings do not negate the utility of this tool as a useful neuropsychological instrument. Nunnally and Bernstein (1993) contend that, in the early stages of construct validation, it is still useful to investigate measures evidencing only modest reliability (that is, 0.70). Furthermore, only after investigating whether corrections for attenuations will increase reliability is it useful to increase the number of items and reduce the measurement error in hopes of enhancing an instrument's reliability. As a point of comparison, the test-retest reliability was 0.60 for the paper-and-pencil MRT. Thus, we interpret our findings to indicate that the VRSR has potential as a reliable measure and will require further study.

2.1.2.3 Concurrent Validity. Pearson Product-Moment correlations between the VRSR time to complete (Speed Index) and all standard measures of neuropsychological functioning yielded a number of statistically significant effects. Note that the direction of all correlations reported below is such that slower VRSR completion time is associated with worse performance on each test. The Speed Index was highly correlated with the Efficiency Index ($r = 0.76$, $p < 0.001$) and moderately with the paper-and-pencil Mental Rotation Test (MRT) ($r = 0.45$, $p < 0.01$). Speed also correlated significantly with tests of visual memory (WMS-Visual Reproduction) under both immediate ($r = 0.50$, $p < 0.006$) and delayed ($r = 0.48$, $p < 0.008$) conditions. There was a significant association with visual attention as measured by Trails A ($r = 0.38$, $p < 0.04$). There were also strong correlations with two measures of executive functioning, one that includes a strong visuoconstructional component (Trails B; $r = 0.46$, $p < 0.01$, WAIS Block Design; $r = 0.64$, $p < 0.001$). Surprisingly, the Speed Index on VR MR testing also was associated with aspects of verbal learning, notably the consistency of word items recalled over the five trials of the California Verbal Learning Test (CVLT) ($r = 0.52$, $p < 0.005$) and the number of perseverations ($r = 0.48$, $p < 0.01$), or times when subjects repeated the same word erroneously. These findings were initially thought to relate to the ability to maintain concentration when presented with a large amount of new information (working memory) than to verbal memory per se. However, in the absence of significant relationships between the Speed Index and Digit Span task, an attention-based hypothesis was not supported. Rather, it may be the case that this finding was the result of shared variance on a general intelligence factor or perhaps more specifically fluid intelligence, as both mental speed tasks and learning (as in the CVLT) have been argued to load on this factor (Peter McGeorge, personal communication, Dec. 12, 2000). Correlations between the Efficiency Index and other neuropsychological tests with the exception of Block Design were not significant.

A comparison of associations between the paper-and-pencil MRT with the other neuropsychological tests provides a useful reference point for interpreting the above correlations. The tests that correlated with the MRT are generally very consistent with the tests that correlated with the VR Speed Index with one notable exception. Although performance on the Judgment of Line Orientation (JLO) was not associated with the VRSR testing, it was strongly correlated with the pencil-and-paper MRT. The JLO is a two-dimensional task that evaluates the ability to perceive spatial orientation. That it would be associated with the ability to mentally rotate two-dimensional portrayals of 3-D objects and not with the ability to physically manipulate 3-D virtual objects may suggest that one of the major cognitive components underlying ability on the MRT relates to the person's ability to construct and manipulate 3-D images from two-dimensional perception. Future investigations in this area are planned.

2.1.2.4 Sex Differences on Rotational Tasks. Men scored significantly better on the MRT given before the VR testing/training ($p < 0.04$). By contrast, there were no differences between men and women on either the Speed Index or the Efficiency Index of the VR testing (p's > 0.8). Interestingly, the difference between men

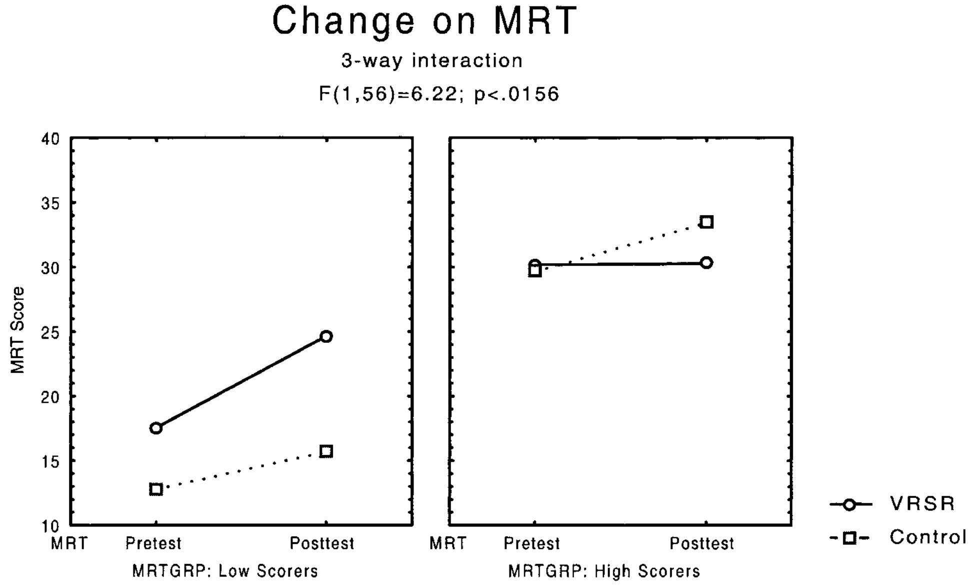

Figure 3. *Pre/post-performance changes on MRT with low/high initial performers for VR versus control groups.*

and women on the MRT after completing the VR training was no longer significant. The existence of gender differences on the MRT is well established, but the mechanism for this difference is not identified. That women can manipulate and successfully rotate 3-D objects as efficiently as men although they cannot visualize the same process as well with 2-D stimuli has potentially useful implications for understanding sex differences in brain functioning that influence visuospatial cognition.

2.1.2.5 Training/Transfer Issues. Subjects showed significant improvement on the VR testing after completing 100 training trials for both the Speed Index ($p < 0.001$) and Efficiency ($p = 0.03$). Subjects in the VR group showed a nonsignificant trend toward improved performance on the MRT ($p < 0.06$). When the changes in MRT performance between the VR and control group were compared by utilizing a split-plot factorial ANOVA, the interaction between group and change over the two testing occasions was nonsignificant. This indicates that VR exposure did not have a specific effect on improving performance among all subjects. However, upon further inspection, this result may be in part due to a low ceiling for MRT performance with the subjects in this sample. Our sample of subjects was notable for performing much better on the MRT than is reported in studies with broader populations. If rotational skills can be trained, it would seem likely that individuals with high existing levels of rotational ability would be less likely to show improvement than would individuals with less ability. In this regard, we examined how individuals who had relatively poor initial MRT scores perform after VR exposure. To directly test this, we divided subjects into groups, based on the MRT scores at the pre-testing. (Note that normative data for this test for comparative purposes can be found in Vandenberg and Kuse (1978).) We used a cutpoint of twenty (out of a possible forty) to create a group of subjects with scores closer to those reported in other studies. Again using a split-plot factorial design, we found a significant ($p < 0.02$) interaction between group (VR and control), MRT group (≤ 20, > 20) and occasion (pre- and post-MRT), such that those low scorers on the MRT who were in the VR group improved significantly more ($F(1,56) = 6.22$; $p < 0.0156$) than other groups. (See figure 3). The VRSR initial-low-performers group produced an average improvement equaling eight points, whereas the control group improvement equaled 2.3

points. This leads to the intriguing suggestion that rotational skills can be trained in VR with initial low performers, and this may lend support for VR rehabilitation strategies aimed at populations with cognitive impairments. Further, VE-delivered, hands-on, spatial rotation training with the MR stimuli may help improve imaginal mental rotation abilities. This assertion is bolstered by a recent study that concluded that rotary object manipulation and mental object rotation share a mutual process that is believed to direct the dynamics of both imagined and actual physical object reorientation (Wohlschlager & Wohlschlager, 1998). Through the use of this VE system with the elderly and persons with TBI or neurological disorders, the feasibility and effectiveness of this novel technology for assessment and rehabilitation purposes will be further explored.

2.2 Study II: Quasi-experimental Follow-up to Study I

Following study I, we conducted a quasi-experimental design investigation to collect pilot data on what factors may have mediated the performance improvement seen on the pencil-and-paper MRT following VRSR training in low initial performers. We videotaped one of the investigators performing well-trained executions of the 140 VRSR trials. The tape was shot in mono-mode (2-D), and the resulting video presented crisp moving images of the blocks being successfully superimposed in a very efficient manner. Twenty-four college students (aged 18 to 36) were administered the pre-test MRT and then asked to watch the videotape, followed by administration of the same alternate form MRT as used in the previous study. In this manner, subjects passively observed the 2-D representations being dynamically superimposed to determine if observation effects might produce equal performance increases as seen with the 3-D VRSR active interaction study (study I). These subjects had initial MRT scores that were similar to those in our "MRT Under-20" sample, and, following simple video observation of 2-D superimposition of block configurations, subject's performance mirrored the control group in the previous study with minimal and nonsignificant gains (two points average gain). Although these data are based on a less controlled, quasi-experimental design approach, they are suggestive that MR training improvements in low initial performers may require the type of active interaction with 3-D stimuli that is characteristic of VRSR training. A more controlled version of this study separately examining active versus passive and 2-D versus 3-D components is currently in the planning stage and will more systematically address this topic.

2.3 Study III: Expanded VR Visuospatial System Experiment with Elderly Subjects—Preliminary Results

Following the completion of study I and II, we expanded our VR visuospatial system by developing a series of 3-D scenarios to investigate depth perception, field dependence (3-D rod and frame test), static/dynamic manual tracking, and visual-field specific reaction time using the ImmersaDesk system. Thirty healthy elderly male and female subjects (aged 65 to 86) were administered these new scenarios along with a shortened version of the VRSR (90 trials instead of 140). The same protocol was followed as in study I, with the additional scenarios administered immediately prior to the VRSR testing. Participants were continuously monitored for any observable signs of discomfort throughout the assessment and encouraged to discontinue participation at any time if ill effects were experienced. Transportation to and from the experimental setting was provided as a precaution against the possibility of impaired driving ability due to potential perceptual aftereffects. The additional visuospatial scenarios included:

Visual Field-Specific Reaction Time Task: This series of tasks tested visual-field specific reaction time performance to stimuli that were presented in a consistent left or right location (regardless of shifts in the subjects' head position) via input from the tracking system contained in the CrystalEyes stereoglasses. Participants were asked to focus their gaze on an "X" presented at the center of the screen and instructed to anticipate a dot to flash either to the left or right of the "X" and to press the response button (using the standard ImmersaDesk response wand) with their thumb as soon as they see the dot. They were told that they will respond best by fixating on the center "X," because they would not be able to anticipant when or from which side the stimulus will appear. Twenty blocks of ten trials were administered (200 trials total) with participants instructed to switch between their left and right hands for alternate blocks of trials. A "water drop" sound was utilized to alert participants to the beginning of a series of ten trials. The stimuli appear at six deg. to the left or right of the midpoint at a width of 0.5 deg. Random intervals are timed from 0.5 to 2.0 sec. and intervals were counterbalanced. The flash duration was at least 0.05 sec., and, if a person missed the flash, the timeout period was 2 sec. after the flash occurs. This task provided a baseline of reaction time performance and data on brain laterality factors that may underlie processing speed for these types of vigilance tasks.

Depth Perception Tasks: The next series of visuospatial tasks consisted of three depth perception scenarios. The first scenario required subjects to match two cubes (identical object depth alignment). Participants were instructed to familiarize themselves to the task by moving the standard ImmersaDesk response wand forwards and backwards in relation to the screen and to notice that this action moved one of the cubes towards or away from them. A static target cube was positioned at varying distances from the subject. The following instructions were given: "The task is to move the cube that you are controlling until it matches the position of the static target cube. When you are satisfied that the cubes are the same distance from you, click the response button of the wand. Respond as accurately as you can. At the tone, the task will begin and when you respond, a chime will sound and you will continue with the next trial of the matching task with the static target cube moving to a new position. Are you ready?" The second depth perception task involved a static target cube and a larger movable ball (dissimilar object depth alignment) that subjects were able to interact with similar to the previous task. The final depth perception task involved two vertical lines (vertical line depth alignment). The participants were instructed to move a line on the right side of the screen by moving the wand back and forth. When the participant judged the two lines to be at the same distance (matched), he/she pushed the response button. After each response, a chime sounded and a new line appeared. Five trials of each depth perception task were administered.

Three-dimensional Field Dependency Task: To assess field dependency, a 3-D virtual rod and frame test was developed. A yellow frame seemingly afloat in space appeared on the screen along with a centered white rod, the orientation of which was controlled by the participant's wand movements. On each trial, the yellow frame was positioned slightly differently from the one previous, and the white bar appeared at a different orientation. Each participant was instructed to respond by pressing the button on the wand when they have positioned the white bar vertically or perpendicular to the floor regardless of the position of the frame. Five trials of this task were presented.

Manual 3-D Tracking Tasks (Static and Dynamic): During the static 3-D manual tracking task, two balls (a blue one on the left and a red one on the right) appeared on the projection screen with a white line running horizontally between them. Participants were instructed to position the wand-controlled crosshairs in the center of the blue ball on the left. They were then instructed to move the ball along the white line using the crosshairs to "push" it, and were told that, to make the ball move, the crosshairs must be in the center of the ball, intersecting the crosshairs closest to the white line. The task involved moving the blue ball to the end of the line where the red ball was, and then back to the original position. The dynamic 3-D manual tracking task required the subject to keep a moving figure ("Tinkerbell") inside a blue 3-D bubble (or orb) that was controlled with the wand. During the first task group, the figure's movement was relatively fluid and stereotypic, consisting of one trial each of x, y, and z rotations (circular paths). The second task group consisted of four paths of different speeds and lengths. During this portion of the task, the figure's actions increased in speed and became increasingly erratic, and the level of difficulty increased for successful tracking. The entire dynamic 3-D manual tracking task took approximately 85 sec. to complete.

The primary purpose of this research was to determine how elderly individuals perform these visuospatial tasks in VR, and then apply this healthy aged group's performance data as a reference sample for future comparisons with elderly persons having various forms of dementia. We were also concerned with measuring the occurrence of VR side effects with this age group and administered pre/post-VR Simulator Sickness Questionnaires (Kennedy et al., 1993) to determine the feasibility of future VR applications with older adults. As well, we plan to compare these results with a younger sample (that is currently being tested) to determine age-related changes in visuospatial performance. In addition to our goal of devising better visuospatial tools for diagnostic purposes, we are acutely interested in determining the additive value of VR for the testing/training of these cognitive processes. Static 2-D tools currently make up the bulk of traditional approaches used in the investigation of visuospatial ability, and it is our view that the dynamic 3-D interactive assets that are available with VR would be better suited to assess and rehabilitate cognitive functions that govern performance in the "real" 3-D world. To this end, comparisons of results with standard visuospatial testing instruments were of considerable interest in this research. How gender affects visuospatial performances was also examined in view of the gender findings on the VRSR in study I. We have begun to analyze this data at the time of this writing and will report later in this article on the analyses that are complete.

2.3.1 Preliminary Results from Elderly Subjects

2.3.1.1 Side Effects. Analyses of the SSQ symptom checklist data have indicated a low occurrence of self-reported negative VR side effects with none of the elderly participants reporting symptoms in the severe range. There were also no individual symptoms that showed a significant increase following VR exposure compared to preexisting self-reported levels. A gender difference was found when the symptoms were analyzed by symptom cluster. The three symptom clusters that are derived from the SSQ (oculomotor, disorientation, and nausea) are produced by combining scores on related individual symptom reports. On the nausea cluster (nausea, stomach awareness, increased salivation, burping, general discomfort, sweating, and difficulty concentrating), women's score indicated less nausea relative to the men pre-VR and more nausea relative to the men following VR exposure ($p = 0.045$).

2.3.1.2 Standard Neuropsychological Test Results. As predicted, there were gender differences, in favor of men, on standard neuropsychological assessment measures of spatial functioning including Judgment of Line Orientation ($t(28) = 5.05$, $p = 0.000$), WAIS-R Block Design ($t(28) = 2.75$, $p = 0.010$), WAIS-R Matrix Reasoning ($t(28) = 2.04$, $p = 0.051$), and WAIS-R Arithmetic ($t(28) = 7.70$, $p = 0.001$). Also, there was a significant difference in favor of men on the paper-and-pencil MRT before ($t(28) = 3.33$, $p = 0.002$) and after training ($t(28) = 2.75$, $p = 0.010$) in the VE.

2.3.1.3 VR Performance. The VRSR reliability was in the high range for internal consistency with this sample. Cronbach's alpha for the "time to complete" measure was 0.96. This value was considerably higher than reliability scores from the younger sample in study I, and further analyses to understand what may underlie this difference are underway. VRSR was also seen to significantly improve when the first twenty testing trials are compared with the final twenty identical trials for both genders ($F(1,28) = 17.05$; $p < 0.0002$). Pre-training mean VRSR time to complete a trial was 11.5 sec. with post-training at 9.4 sec. However, in contrast to study I, males significantly outperformed females at both VRSR times ($F(1,28) = 6.62$; $p < 0.016$). The mean time to complete a VRSR trial for males was 7.9 sec., with the female mean time at 13.1/sec. By contrast, on the one other VR-delivered visuospatial task that we have analyzed thus far, the similar object depth perception task, no gender differences in performance were found. Analyses of the remaining depth perception, field dependence (3-D rod and frame test), static/dynamic manual tracking, and visual-field specific reaction time tasks are being conducted at the time of this writing.

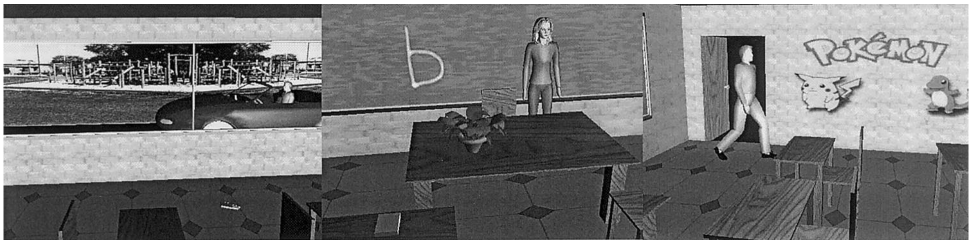

Figure 4. *Scenes from the virtual classroom for assessment of Attention Deficit Hyperactivity Disorder.*

Three directions will be pursued in the ongoing data analyses.

1. Comparisons between standard neuropsychological test results and VR performance testing to determine how a VR-based, interactive 3-D approach to molecular visuospatial assessment differs from traditional measures and how this may better address aging and gender-based performance.
2. Determine the relative components of spatial rotation performance by analyzing the hypothesized subcomponents of speed of processing, depth perception, field dependency, and static/dynamic manual tracking skill. These analyses will utilize the full series of VR tasks as covariants to determine how these subcomponents individually contribute to the variability seen in both pencil-and-paper and VR measures of mental/spatial rotation ability.
3. Creation of a database to be used to compare performance with populations with CNS dysfunction (such as Alzheimer's, Stroke, and TBI) and younger participants on the VR tasks.

Following this research, the longer-term goal is to develop a 3-D desktop system using CrystalEyes shutter glasses to run comparative tests to determine if these scenarios can be successfully delivered on a less expensive and more accessible platform. The aim is to produce a suite of standardized VE-delivered 3-D visuospatial assessment and rehabilitation tools, and we are currently designing new tasks to target 3-D line bisection, figure/ground judgments, and other perceptual reasoning targets that may be of use by both researchers and clinicians. More-detailed information on the rationale, equipment, and methodology for this work can be found in McGee et al. (2000).

3 Cognitive Process Applications Using HMDs

A second line of research that is being addressed in our lab concerns the development of a series of ecologically valid functional VE scenarios to serve as platforms for addressing a range of cognitive and functional processes. Our first effort in this direction has been in the development of a HMD-delivered classroom scenario for the assessment and rehabilitation of attention processes. (See figure 4).

Although this platform is ultimately envisioned to be capable of delivering cognitive testing and training protocols that could address other cognitive processes (memory and executive functions), we are initially targeting attention. Attention processes are the gateway to information acquisition and serve as a necessary foundation for most higher learning. Impairments in attention can be found in clinical populations across the lifespan and are commonly observed in persons diagnosed with Attention Deficit Hyperactivity Disorder (ADHD), TBI, and as a feature of various forms of age related dementia (such as Alzheimer's disease). Little clinically based VR work has been done with this "basic" gateway cognitive process thus far, which is surprising in view of the relative significance of attention impairments seen in a variety of clinical conditions. More-effective assessment and rehabilitation tools are needed to address attention processes for a variety of reasons. In children, attention skills are the necessary foundation for future educational activities. Specific to ADHD, improved assessment of attention is vital for diagnostic purposes, special education placement decisions, determination of the use and effectiveness of pharmacological treatments, and for outcome measurement following interventions. Regarding TBI, even with mild trauma, these patients often suffer attention deficits that require focus as a precursor to rehabilitative work on higher cognitive processes (that is, memory, executive functions, and problem solving). Also, even if higher processes are unable to be remediated in cases of severe TBI, some level of attention ability is essential for vocational endeavors, functional independence, and quality-of-life pursuits. With the elderly, a more fine-grained assessment of basic attention deficits may provide an early indicator of dementia-related symptoms, could suggest functional areas in which an older person might be at risk (that is, automobile driving, operating machinery, and so forth), and guide development of compensatory strategies useful to maximize or maintain functional independence. HMDs are well suited for these types of applications because they present a controlled stimulus environment in which cognitive/attention challenges can be administered along with the precise control of “distracting” auditory and visual stimuli. This level of experimental control may also allow for the development of attention assessment and rehabilitation tasks more similar to what is found in the real world.

Our first effort in this area has involved the development of a virtual “classroom” specifically aimed at the assessment of ADHD. VE technology appears to provide specific assets for addressing impairments seen in ADHD that are not available using existing methods. The scenario consists of a standard rectangular classroom environment containing desks, a male or female teacher, a blackboard across the front wall, a side wall with a large window looking out onto a playground and street with vehicles and people, and, on each end of the opposite wall, a pair of doorways through which activity occurs. Within this scenario, children’s attention performance is assessed while a series of typical classroom distracters (that is, immersive audio-supported ambient classroom noise, movement of other pupils, activity occurring outside the window, and so on) are systematically controlled and manipulated within the virtual environment. The child sits at a real desk while seeing a virtual desk in the HMD within the virtual classroom. On-task attention can be measured in terms of performance (reaction time) on a variety of attention challenges that can be adjusted based on the child’s expected age/grade level of performance. For example, on the simpler end of the continuum, the child can be required to press a response button upon the direct instruction of the virtual teacher or whenever the child hears the name of a specific target color mentioned by the teacher (focused or selective attention task). Sustained attention can be assessed by manipulating the time demands of the testing. More-complex demands requiring alternating or divided attention can be developed whereby the student needs to respond only when the teacher states the target color in relation to an animal (the brown dog, as opposed to the statement, “I like the color brown”) and only when the word *dog* is written, or a picture of a dog appears on the blackboard. In addition to attention-driven reaction time performance, behavioral measures that are correlated with distractibility and/or hyperactivity components (head turning, gross motor movement), and impulsive nontask behaviors (time playing with “distracter” items on the desk) can be measured via strategically located trackers. Our first study is presently comparing ADHD diagnosed children (aged eight to twelve) with a nondiagnosed control group using more-basic attention challenges that are commonly seen on currently available continuous performance tasks and in common classroom tasks (listen-look-respond).

This work is currently in progress and is in the user-centered design phase. In this phase, we are testing children on basic selective and alternating attention tasks, soliciting their feedback pertaining to aesthetics and usability of the VE, and incorporating some of their comments into the actual iterative design-evaluate-redesign cycle. Thus far, we have tested fourteen nondiagnosed children (aged six through twelve), and initial results indicate little difficulty in adapting to use of the HMD, no self-reported occurrence of side effects deter-

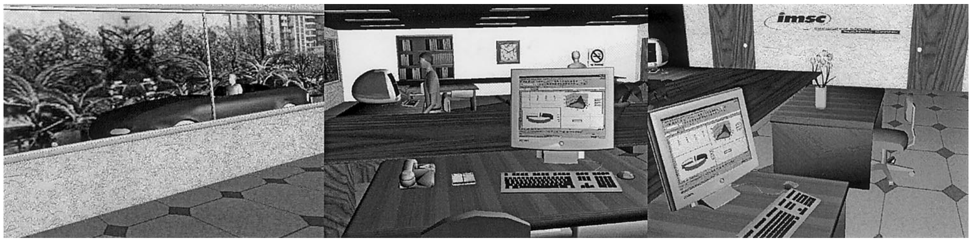

Figure 5. *The virtual office scenario.*

mined by post-interviews using the Simulator Sickness Questionnaire (Kennedy et al., 1993), and excellent performance on the stimulus-tracking challenges. Our initial clinical trial is scheduled to commence in the fall of 2000. More detailed information on the rationale, equipment, and methodology for this project can be found in Rizzo et al. (2000).

Other scenarios (such as work situations, and home environments) using the same logic and approach are being conceptualized and developed to address cognitive/functional processes that are relevant for a range of other clinical populations. For example, we have now constructed a virtual office environment that evolved from expanding some of the basic design elements of the classroom VE. (See figure 5).

As with the classroom VE, the user will sit at a real desk, but, within the HMD, they will see the scenes that compose the virtual office. The virtual desk contains a phone, computer screen, and message pad, while, throughout the office, a virtual clock ticks in realtime, and a variety of human avatar representations of coworkers/supervisors can be actively engaged. Various performance challenges can be delivered via the computer screen (visual mode), the phone (auditory mode), and from the avatar supervisors' verbal directions. These commands can direct the user to perform certain functions within the environment that can be designed to assess and rehabilitate attention, memory, and executive functions. For example, to produce "prospective" memory challenges, the user might receive a command from the virtual supervisor to turn on the computer at a specific time to retrieve a message that will direct a response. This will require the user to hold this information in mind, monitor the time via the wall clock, and then initiate a response at the appropriate time. By adding multiple contingent commands, both attention and executive functioning can be addressed. As well, the influence of distraction can be tested or trained for via the presentation of ambient office sounds (such as radio announcements, conversations, and so on), activity occurring outside the window (cars rumbling by), or by producing extraneous stimuli on the desktop (irrelevant, yet attention-grabbing email messages appearing on computer screen). Essentially, functional work performance challenges typical of what occurs in the real world can be systematically presented in an ecologically valid VE.

4 Conclusions

In closing, the complexity of these scenarios and their potential usefulness are primarily governed by the developer's technical skill, imagination, and an informed assessment of the needs, assets, and limitations of the intended users. The design and implementation of such virtual environments for persons with cognitive/functional limitations illustrate a unique set of challenges and opportunities that will require interdisciplinary cooperation between usability specialists and scientists who have domain-specific knowledge of the range of impairments seen with clinical populations. User-centered design and evaluation methods are essential to this process throughout the application lifecycle. If successful, VR technology could serve to revolutionize the standard approaches used for neuropsychological assessment and cognitive rehabilitation, while at the same time promoting the development of more-usable and-accessible systems for unimpaired populations.

References

Army Individual Test Battery. (1944). *Manual of Directions and Scoring.* Washington, D.C.: War Department, Adjutant General's Office.

Benton, A. L., Varney, N. R., & Hansher, K. S. (1978). Visuospatial judgment: A clinical test. *Archives of Neurology, 36*, 364–367.

Brown, D. J., Kerr, S., & Wilson, J. R. (1997). Virtual environments in special-needs education. *Comm. of the ACM, 40*(8), 72–75.

Casey, M. B., Nuttall, R., Pezaris E., & Benbow, C. P. (1995). The influence of spatial ability on gender differences in mathematics college entrance test scores across diverse samples. *Developmental Psychology, 31*, 697–705.

Cole, K. C. (1998, March 14). Hawking's universe is open and shut. *Los Angeles Times,* pp. A1, A19.

Delgado, A., & Prieto, G. (1996). Sex differences in visuospatial ability: Do performance factors play such an important role? *Memory and Cognition, 24*(4), 504–510.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. (1983). *California Verbal Learning Test.* New York: Psychological Corporation.

Gainotti, G., Parlato, V., Monteleone, D., & Carlomagno, S. (1992). Neuropsychological markers of dementia on visual spatial tasks: A comparison between Alzheimer's type and vascular forms of dementia. *Journal of Clinical and Experimental Neuropsychology, 14,* 239–252.

Gouchie, C., & Kimura, D. (1991). The relationship between testosterone levels and cognitive ability patterns. *Psychoneuroendocrinology, 16,* 323–324.

Kampen, D. L., & Sherwin, B. B. (1994). Estrogen use and verbal memory in healthy postmenopausal women. *Obstetrics and Gynecology, 83,* 979–983.

Kennedy, R. S., & McCauley, M. E. (1984). *The Motion History Questionnaire.* Orlando: Essex Corporation.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness. *International Journal of Aviation Psychology, 3*(3) 203–220.

Lezak, M. D. (1995). *Neuropsychological assessment.* New York: Oxford University Press.

McGee, J. S., van der Zaag, C., Rizzo, A., Buckwalter, J. G., Neumann, U., & Thiebaux, M. (2000). Issues for the assessment of visuospatial skills in older adults using virtual environment technology. *CyberPsychology and Behavior, 3*(3), 469–482.

McGeorge, P. (2000). Personal communication, Dec. 12, 2000.

Masters, M. S., & Sanders, B. (1993). Is the gender difference in mental rotation disappearing? *Behavior Genetics, 23*(4), 337–341.

Miller, B. L., Ikonte, C., Ponton, M., Levy, M., Boone, K., Darby, A., Berman, N., Mena, I., & Cummings, J. L. (1997). A study of the Lund Manchester research criteria for frontotemporal dementia: Clinical and single-photon emission CT correlations. *Neurology, 48,* 937–942.

Mitrushina, M. N., Boone, K. B., & D'Elia, L. F. (1999). *Handbook of normative data for neuropsychological assessment.* New York: Oxford University Press.

National Institute of Disability and Rehabilitation Research. *Notice of final long-range plan for fiscal years 1999–2004.* [On-line] Available: *www.ncddr.org/rerc/issues.html.*

National Science Foundation. *Program announcement (NSF PA# 98-168)* [On-line] Available: *www.nsf.gov/pubs/1999/nsf98168/nsf98168.htm.*

Newcombe, N., Bandura, M. M., & Taylor, B. G. (1983). Sex differences in spatial ability and spatial activities. *Sex Roles, 9,* 530–539.

Nunnally, J. C., & Bernstein, I. (1993). *Psychometric theory.* New York: McGraw-Hill.

Peters, M., Laeng, B., Lathan, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn Vandenberg and Kuse Mental Rotations Test: Different versions and factors that affect performance. *Brain and Cognition, 28,* 39–58.

Pugnetti, L., Mendozzi, L., Motta, A., Cattaneo, A., Barbieri, E., & Brancotti, S. (1995). Evaluation and retraining of adults' cognitive impairments: Which role for virtual reality technology? *Computers in Bio. and Medicine, 25*(2), 213–227.

Qubeck, W. J. (1997). Mean differences among subcomponents of Vandenberg's Mental Rotation Test. *Perceptual and Motor Skills, 85*(1), 323–332.

Richardson, J. T. (1994). Gender differences in mental rotation. *Perceptual and Motor Skills, 78*(2), 435–448.

Rizzo, A., & Buckwalter, J. G. (1997). Virtual reality and cognitive assessment and rehabilitation: The state of the art. In

G Riva (Ed.), *Virtual reality in neuro-psycho-physiology: Cognitive, clinical, and methodological issues in assessment and rehabilitation* (pp. 123–146) Amsterdam: IOS Press.

Rizzo, A., Buckwalter, J. G., Humphrey, L., van der Zaag, C., Bowerly, T., Chua, C., Neumann, U., Kyriakakis, C., van Rooyan, A., & Sisemore, D. (2000). The virtual classroom: A virtual environment for the assessment and rehabilitation of attention deficits. *CyberPsychology and Behavior, 3*(3), 483–499.

Rizzo, A., Buckwalter, J. G., & van der Zaag, C. (in press). Virtual environment applications in clinical neuropsychology. In K. Stanney (Ed.), *Handbook of Virtual Environments,* New York: Erlbaum Publishing.

Rose, F. D. (1996). Virtual reality in rehabilitation following traumatic brain injury. *Proceedings of the European Conference on Disability, Virtual Reality and Associated Technology,* 5–12.

Sattler, J. (1988). *Assessment of children*. San Diego: J. Sattler Publishers.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects, *Science, 171,* 701–703.

Signorino, M., Pucci, E., Brizioli, E., Cacchio, G., Nolfe, G., & Angeleri, F. (1996). EEG power spectrum typical of vascular dementia in a subgroup of Alzheimer patients. *Archives of Gerontology and Geriatrics, 23,* 139–151.

Silverman, I., & Phillips, K. (1993). Effects of estrogen changes during the menstrual cycle on spatial performance. *Ethology and Sociobiology, 14,* 257–270.

Stephanidis, C., Salvendi, G., Akoumianakis, N., Bevan, N., Brewer, J., Emiliani, P. L., Galetsas, A., Haataja, S., Lakovidis, I., Jacko, J. A., Jenkins, P., Karshmer, A. I., Korn, P., Marcus, A., Murphy, H. J., Stary, S., Vanderherden, G., Weber, G., & Ziegler, J. (1998). Toward an information society for all: An international research and development agenda. *International Journal of Human-Computer Interaction 10*(2), 107–134.

Vandenberg, S. G., & Kuse, A. R., (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills* 47, 599–604.

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. *Psychological Bulletin* 117, 250–270.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised*. New York: The Psychological Corporation.

———. (1987). *Wechsler Memory Scale—Revised*. New York: The Psychological Corporation.

Wohlschlager, A., & Wohlschlager, A. (1998). Mental and manual rotation. *Journal of Experimental Psychology* 24, 397–412.