Effects of Handling Real Objects and Self-Avatar Fidelity on Cognitive Task Performance and Sense of Presence in Virtual Environments

Benjamin Lok, University of Florida
Samir Naik, Disney Imagineering
Mary Whitton and Frederick Brooks, UNC-Chapel Hill

Abstract

Immersive virtual environments (VEs) provide participants with computer-generated environments filled with virtual objects to assist in learning, training, and practicing dangerous and/or expensive tasks. But does having every object being virtual inhibit the interactivity and level of immersion? If participants spend most of their time and cognitive load on learning and adapting to interacting with virtual objects, does this reduce effectiveness of the VE?

We conducted a study that investigated how handling real objects and self-avatar visual fidelity affects performance and sense-of-presence on a spatial cognitive manual task. We compared participants' performance of a block arrangement task in both a real-space environment and several virtual and hybrid environments. The results showed that manipulating real objects in a VE brings task performance closer to that of real space, compared to manipulating virtual objects. There was not a significant difference in reported sense-of-presence, regardless of the self-avatar's visual fidelity or the presence of real objects.

Keywords: Virtual Environments, Sense of Presence, Human-Computer Interaction

1. Introduction

1.1. Motivation

Conducting design evaluation and assembly feasibility evaluation tasks in immersive virtual environments (VEs) has become one of the major productive applications of VEs [1]. The ideal VE system would have the participant believe that he was actually performing a task. Parts and tools would have mass, feel real, and handle appropriately. The participant would naturally interact with the virtual world, and in turn, the virtual objects would respond to the participant's action appropriately [2].

1.2. Current VE Methods

Obviously, current VEs are far from that ideal system. Indeed, not interacting with every object as if it were real has distinct advantages, as in dangerous or expensive tasks. In current VEs, almost all objects in the environment are virtual, but both assembly and servicing are hands-on tasks, and the principal drawback of virtual models — that there is nothing there to feel, nothing to give manual affordances, and nothing to constrain motions — is a serious one for these applications. Simulating a wrench with a six degree-of-freedom wand, for example, is far from realistic, perhaps too unrealistic to be useful.

Interacting with purely virtual objects could impose three limiting factors:

- Limits the types of feedback, such as motion constraints and haptics, the system could provide the user.

- The VE representation of real objects (real-object avatars) is usually stylized and not necessarily visually faithful to the object itself.

- Hinders real objects (including the user) from naturally interacting with virtual objects.

This work investigates the impact of these factors on task performance and sense of presence in a spatial cognitive task. Most design verification and training tasks are cognitive.

In this work, we extend our definition of an *avatar* to include a virtual representation of any real object, not just the participant. The *real-object avatar* is registered with the real object, and ideally, they are faithful in look, form, and function with the real object. The *self-avatar* refers specifically to the user's virtual representation.

We believe a *hybrid environment* system – a system that handles dynamic real objects – would be effective in providing natural interactivity and visually-faithful self-avatars. In turn, this should improve task performance and sense of presence.

2. Previous Work

2.1. Self-Avatars

The user is represented within the VE by a self-avatar chosen from a library of representations, a generic self-avatar, or no self-avatar. A survey of VE research shows the most common approach is a generic self-avatar – literally, one size fits all [1]. The participant's self-avatars are typically stylized human models, such as those found in commercial packages. While containing a substantial amount of detail, these models do not visually match a participant's appearance.

Studies have shown that providing generic self-avatars substantially improves sense-of-presence over providing no self-avatar [3]. However, researchers hypothesize that the visual misrepresentation of self would reduce how much a participant believed he was "in" the virtual world, his *sense-of-presence*. Usoh hypothesizes, "Substantial potential presence gains can be had from tracking all limbs and customizing [self-]avatar appearance [4]."

Recent studies suggest that even crude self-avatar representations convey substantial information for navigation, social interaction, and task performance [5]. With self-avatars, emotions such as embarrassment, irritation, and self-awareness could be generated [6][7].

Providing realistic self-avatars requires capturing the participant's motion, shape, and appearance. In general, VE systems attach extra trackers to the participant for sensing changing positions to drive an articulated stock self-avatar model. Presenting and controlling an accurate representation of the participant's shape and pose is complicated by the human body's deformability and numerous degrees of freedom. Dynamically matching the virtual look to the physical reality is difficult, though there are commercial systems, such as the AvatarMe system, that generate static-textured, personalized self-avatars [8].

2.2. Interactions in VEs

Ideally, a participant should be able to interact with the VE by natural speech and natural body motions. The VE system would understand and react to expressions, gestures, and motion. The difficulty is in capturing this information for simulation input.

12/10/2003

The fundamental interaction problem is that most things are not real in a VE. In effort to address this, some VEs provide tracked, instrumented real objects as input devices. Common interaction devices include an articulated glove with gesture recognition or buttons (Immersion's Cyberglove), tracked mouse (Ascension Technology's 6D Mouse), or tracked joystick (Fakespace's NeoWand).

Another approach is to engineer a specific device for interaction. This typically improves interaction affordance, such that the participant interacts with the system in a more natural manner. For example, Hinckley, *et al.*, used a tracked doll's head with props to more naturally select cutting planes for visualizing MRI data [9]. However, this specialized engineering is time-consuming and often usable for only a particular type of task. VE interaction studies have been done on interaction ontologies [10], interaction methodologies [11], and 3-D GUI widgets and physical interaction [12].

3. User Study

3.1. Study Goals

We investigated the following questions: For cognitive tasks,

- Does ***interacting with real objects*** and ***seeing a visually faithful self-avatar*** improve task performance?

- Does ***seeing a visually faithful self-avatar*** improve sense-of-presence?

We employed a system that can incorporate dynamic real objects into a VE. It uses multiple cameras to generate virtual representations of real objects at interactive rates [13]. Thus we

could investigate how cognitive tasks performance is affected by interacting with real versus virtual objects. The results would be useful for training and assembly verification VEs, which often require problem solving while interacting with tools and parts.

Video capture of real object appearance also has another potential advantage — enhanced visual realism. Generating virtual representations of the participant in real time would allow the system to render a visually faithful self-avatar. The real-object appearance is captured from a camera that has a similar line of sight as the participant. The system allows us to investigate whether having a visually faithful self-avatar, as opposed to a generic self-avatar, increases sense-of-presence and task performance. This will be useful for immersive virtual environments that aim for high sense-of-presence, such as phobia treatment and entertainment VEs.

3.2. Task Description

We sought to abstract tasks common to VE design applications. We specifically wanted to use a task that focused on cognition and manipulation over participant dexterity or reaction speed because of current technology, typical VE applications, and physical variability among participants. We conducted a user study on a block arrangement task. We compared a purely virtual task system and two hybrid task systems that differed in level of visual fidelity. In all three cases, we used real-space performance as a baseline.

The task we designed is similar to, and based on, the block design portion of the Wechsler Adult Intelligence Scale (WAIS). Developed in 1939, the Wechsler Adult Intelligence Scale is a test

widely used to measure IQ [14]. The block-design component measures reasoning, problem solving, and spatial visualization.

In the standard WAIS block design task, participants manipulate one-inch cubes to match target patterns. We modified the task to still require cognitive and problem solving skills while focusing on interaction methodologies. The small one-inch cubes of the WAIS would be difficult to manipulate with traditional VR approaches and hamper the reconstruction system due to reconstruction error. We used three-inch cubes, as shown in Figure 1.
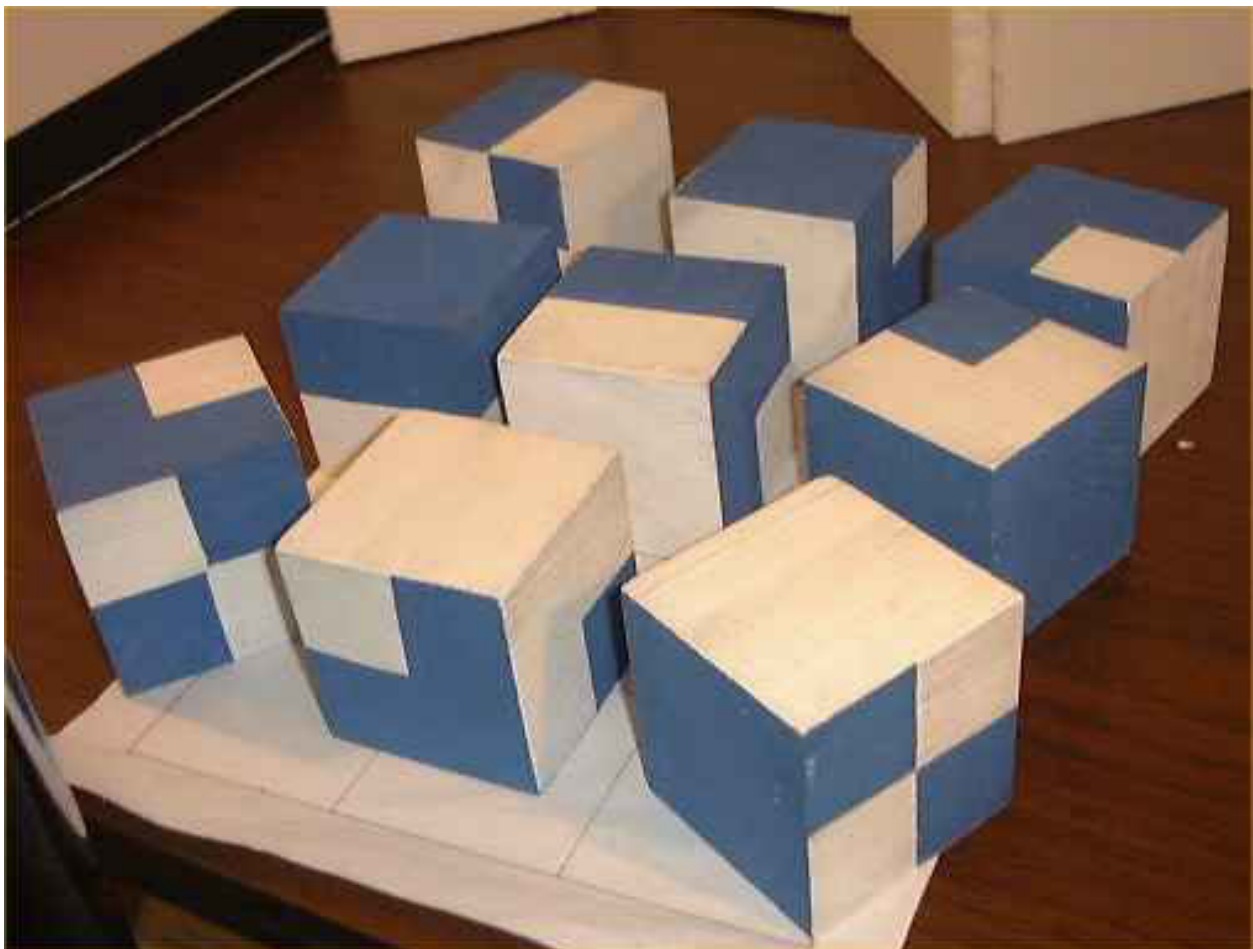
Participants manipulated four or nine identical wooden blocks to make the top face of the blocks match a target pattern. Each cube had six patterns on its faces that represented the possible quadrant-divided white-blue patterns. There were two target patterns sizes, *small* four-block patterns in a 2 x 2 arrangement, and *large* nine-block patterns in a 3 x 3 arrangement.

3.3. Task Design

The user study was a between-subjects design. Each participant performed the task in a real space environment (RSE), and then one of three VE conditions (Figure 2). The independent variables were the VE interaction modality (real or virtual blocks) and the VE self-avatar visual fidelity (generic or visually faithful). The three VE conditions had:

- Virtual objects, generic self-avatar (purely virtual environment - PVE)

- Real objects, generic self-avatar (hybrid environment - HE)

- Real objects, visually faithful self-avatar (visually-faithful hybrid environment - VFHE)

The participants were randomly assigned to one of the three groups, 1) RSE then PVE, 2) RSE then HE, or 3) RSE then VFHE.
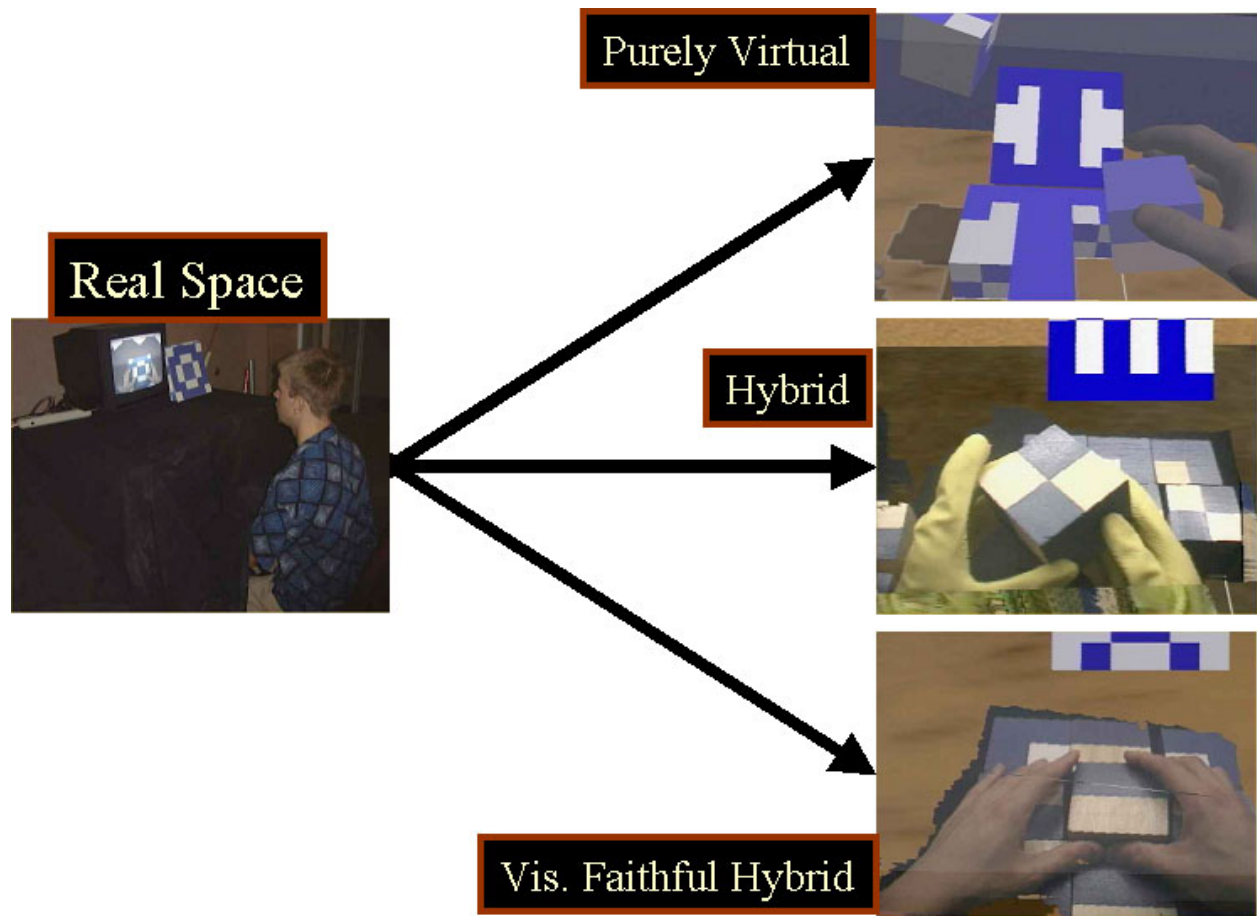
The task was accessible to all participants, and the target patterns were intentionally of a medium difficulty (determined through pilot testing). Our goal was to use target patterns that were not so cognitively easy as to be manual dexterity tests, nor so difficult that participant spatial ability dominated the data.

Real Space Environment (RSE). The participant sat at a desk (Figure 3) with a rectangular enclosure. The enclosure was draped with a dark cloth and the enclosure side facing the participant was open. A television atop the enclosure displayed the video feed from a camera mounted inside the enclosure. The camera had a similar line of sight as the participant, and the participant performed the task while watching the TV.
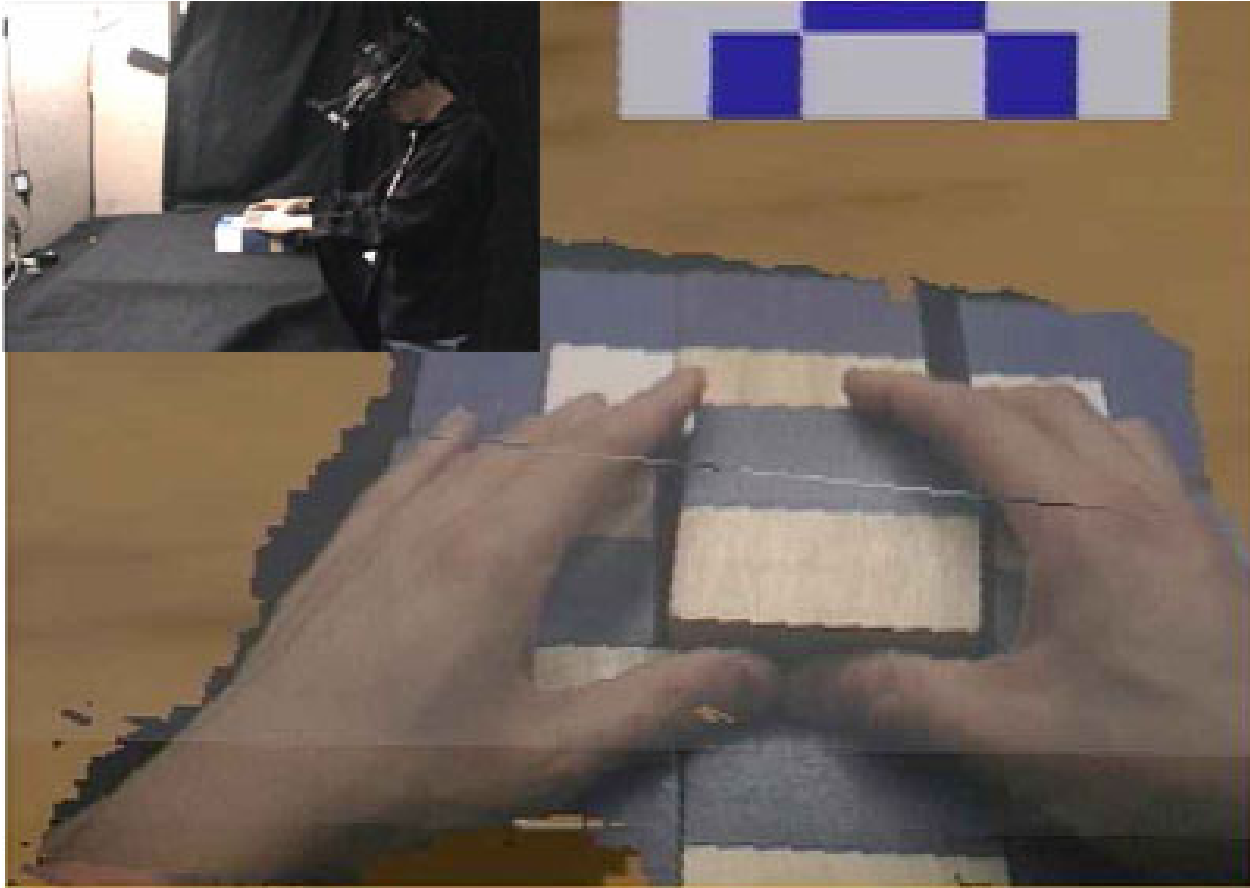
Purely Virtual Environment (PVE). Participants stood at a four-foot high table, and wore Fakespace Pinchgloves, tracked with Polhemus Fastrak trackers, and a Virtual Research V8 head-mounted display (HMD) (Figure 4). The participant picked up a virtual block by pinching two fingers together (i.e. thumb and forefinger). When the participant released the pinch, the virtual block was dropped and an open hand avatar was displayed. The self-avatar's appearance was generic (a neutral gray color).

The block closest to an avatar's hand was highlighted to identify which block would be selected by pinching. Pinching caused the virtual block to snap into the avatar's hand. To rotate the block, the participant rotated his hand while maintaining the pinching gesture.

Releasing the block within six inches of the virtual table caused it to snap into an unoccupied position in a three by three grid on the table. This reduced the need for fine-grained interaction that would have inflated the time to complete the task. Releasing the block away from the grid caused it to drop onto the table. Releasing the block more than six inches above the table caused the block to float in mid-air to aid in rotation. There was no inter-block collision detection, and block interpenetration was not automatically resolved.
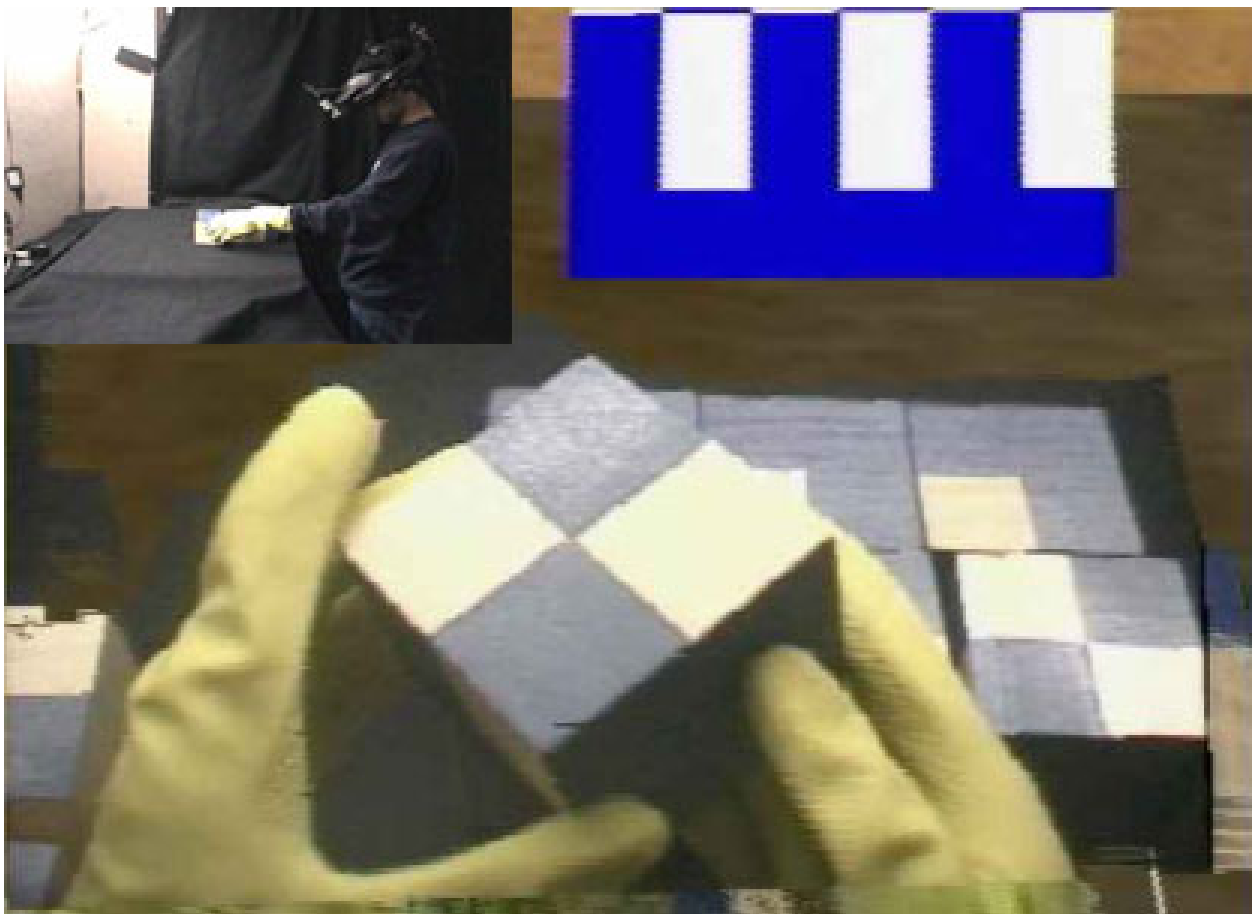
Hybrid Environment (HE). Participants wore yellow dishwashing gloves and the HMD (Figure 5). Within the VE, participants handled physical blocks, identical to the RSE blocks, and saw a self-avatar with accurate shape and generic appearance (due to the gloves).

Visually-Faithful Hybrid Environment (VFHE). Participants wore only the HMD. The self-avatar was visually faithful, as the reconstruction of the user's hands was texture-mapped with images from a HMD mounted camera (Figure 6).

Virtual Environment. The VE room was identical in all three of the virtual conditions (PVE, HE, VFHE). It had several virtual objects, including a lamp, plant, and painting, along with a virtual table that was registered with a real Styrofoam table. The enclosure in the RSE was also rendered with transparency in the VE (Figure 7).

All the VE conditions were rendered on an SGI Reality Monster. The PVE ran on one rendering pipe at a minimum of twenty FPS. The HE and VFHE ran on four rendering pipes at a minimum of twenty FPS for virtual objects and twelve FPS for reconstructing real objects. The reconstruction system used 4 cameras, with 0.3 seconds of estimated latency and 1 cm reconstruction error. The participant wore a HMD (640 x 480 resolution) that was tracked with a 3rdTech HiBall optical tracker.

Rationale for Conditions. We expected a participant's real space environment performance would produce the best results due to the natural interaction and good visual fidelity. Thus, we

compared how close a participant's task performance in VE was to their RSE task performance. We compared the reported sense-of-presence in the VE conditions to each other.

In a pilot study (*n*=20), participants performed the RSE task on a table without the enclosure and monitor. There was no difference in task performance compared to the RSE enclosure and monitor setup. The enclosure and camera allowed the RSE to have a similar field of view and working volume as the VE conditions.

The RSE was used for task training to reduce variability in individual task performance and as a baseline. The block design task had a learning curve (examined through pilot testing), and performing the task in the RSE allowed participants to practice without spending additional time in the VE (limited to fifteen minutes – determined through pilot testing).

- The PVE was a *plausible VE approach* to the block task. As in current VEs, most of the objects were virtual, and interactions were done with specialized equipment and gestures. The difference in task performance between the RSE and PVE corresponded primarily to the impedance of interacting with virtual objects.

- The HE evaluated the effect of real objects on task performance. We assumed any interaction hindrances caused by the gloves were minor compared to the effect of handling real objects.

- The VFHE evaluated the cumulative effect of both real object interaction and visually faithful self-avatars on performance.

3.4. Measures

Task Performance. Participants were timed on replicating *correctly* the target pattern. We recorded if the participant incorrectly concluded that target pattern was replicated. In these cases, the participant was informed and continued to work on the pattern until correct.

Sense-of-presence. Participants answered the Steed-Usoh-Slater Presence Questionnaire (SUS) [15].

Other Factors. Participants answered the Guilford-Zimmerman Aptitude Survey, Part 5: Spatial Orientation (spatial ability) and the Kennedy – Lane Simulator Sickness Questionnaire (simulator sickness). Participants were interviewed after completing the task. We recorded participant- and experimenter-reported behaviors.

3.5. Experiment Procedure

All participants completed a consent form and questionnaires to gauge their physical and mental condition, simulator sickness, and spatial ability.

Real Space. Next, the participant performed the task in real space environment (RSE). The participant examined the blocks, the cloth on the enclosure was lowered, and the TV turned on. The participant did six practice patterns, three small (2 x 2) and then three large (3 x 3). The participant was told the number of blocks involved for a pattern, and to notify the experimenter when they were done. Next, the participants did six timed test patterns, three small and three

large. Between patterns, the participant was asked to randomize the blocks' orientations. The order of patterns for each participant was unique.

Virtual Space. Next, the participant entered a different room where the experimenter helped the participant put on the HMD and any additional VE equipment (PVE – tracked pinch gloves, HE – dishwashing gloves). Following a period of VE adaptation, the participant practiced on two small and two large patterns. The participant then was timed on two small and two large test patterns. A participant could ask questions and take breaks between patterns if they desired. Only one person (a PVE participant) asked for a break.

After the VE, the participant was interviewed and filled out questionnaires.

Managing Anomalies. If the head or hand tracker lost tracking or crashed, we quickly restarted the system (about 5 seconds). In almost all the cases, the participants were so engrossed with the task they did not even *notice* the lack of tracking. We noted long or repeated tracking failures, and participants who were tall were allowed to sit to perform the task. On hand were additional patterns for replacement of voided trials, such as if a participant dropped a block onto the floor. None of these anomalies appeared to significantly affect task performance.

3.6. Hypotheses

Task Performance. Participants who ***manipulate real objects*** in the VE (HE, VFHE) will complete the spatial cognitive manual task significantly closer to their RSE task performance than will participants who ***manipulate virtual objects*** (PVE), i.e. *interacting with real objects*

improves task performance. Further, there will not be a significant difference in task performance for VFHE and HE participants, i.e. *the presence of real objects would have the similar effects on sense of presence regardless of self-avatar fidelity*.

Sense-of-Presence. Participants represented in the VE by a ***visually faithful self-avatar*** (VFHE) will report a higher sense-of-presence than will participants represented by a ***generic self-avatar*** (PVE, HE), i.e. *avatar visual fidelity increases sense-of-presence*. Further, there will not be a significant difference in sense-of-presence for HE and PVE participants, i.e. *generic self-avatars would have similar effects on sense-of-presence regardless of the presence of real objects*.
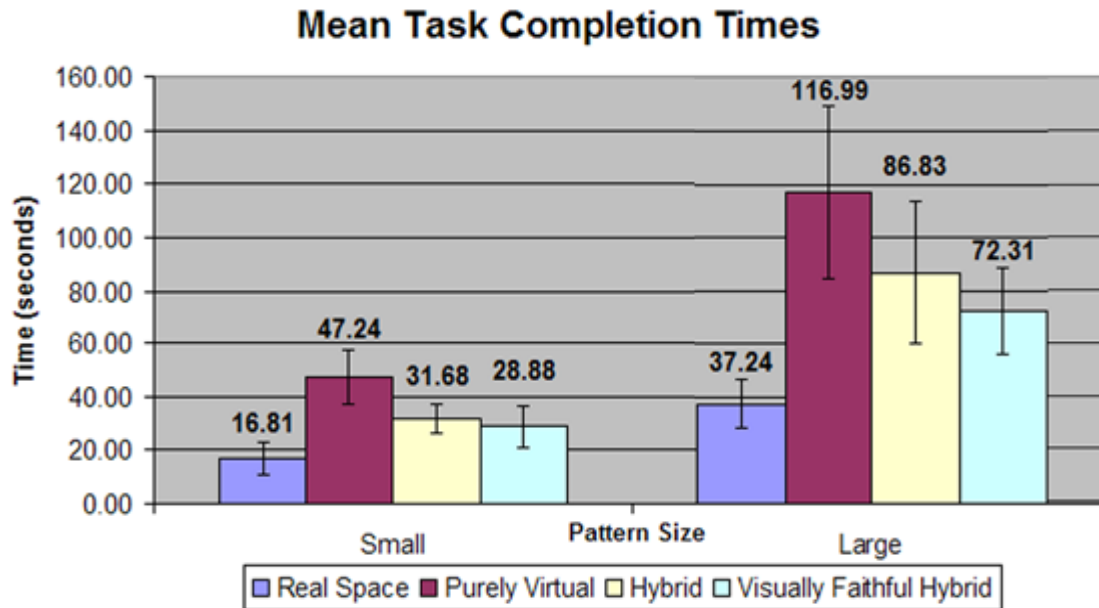
4. Results

4.1. Subject Information

Forty participants completed the study, thirteen in the purely virtual environment (PVE) and hybrid environment (HE), and fourteen in the visually-faithful hybrid environment (VFHE). Participants were mostly male (thirty-three) undergraduate students at UNC-CH (thirty-one). They were primarily recruited from UNC-CH Computer Science classes and word of mouth.

They reported little prior VE experience (M=1.37, s.d.=0.66), high computer usage (M=6.39, s.d.=1.14), and moderate – 1 to 5 hours a week – computer/video game play, on [1..7] scales. There were no significant differences between the groups. We required participants to have taken or be currently enrolled in a higher-level mathematics course (*i.e.* Calculus 1). This reduced participant spatial ability variability, and in turn reduced task performance variability.

4.2. Task Performance

The dependent variable for task performance was the difference in the time to correctly replicate the target pattern in the VE condition compared to the RSE.

Table 1 – Task performance results

	Small Pattern Time (seconds)		Large Pattern Time (seconds)	
	Mean	S.D.	Mean	S.D.
RSE (n=40)	16.81	6.34	37.24	8.99
PVE (n=13)	47.24	10.43	116.99	32.25
HE (n=13)	31.68	5.65	86.83	26.80
VFHE (n=14)	28.88	7.64	72.31	16.41

Table 2 – Difference between VE and RSE times

	Small Pattern Time (seconds)		Large Pattern Time (seconds)	
	Mean	S.D.	Mean	S.D.
PVE – RSE	28.28	13.71	78.06	28.39
HE – RSE	15.99	6.37	52.23	24.80
VFHE – RSE	13.14	8.09	35.20	18.03

Table 3 – Between groups task performance

	Small Pattern		Large Pattern	
	t-test	p-value	t-test	p-value
PVE - RSE vs. VFHE - RSE	3.62^1	0.0013^{**}	4.72^1	0.0001^{***}
PVE - RSE vs. HE - RSE	3.12^1	0.0047^{**}	2.77^1	0.011^*
VFHE – RSE vs. HE - RSE	1.01^2	0.32^+	2.05^2	0.055^+

[1]One-tailed t-test with unequal variances and significant at [*]α=0.05, [**]α=0.01, [***]α=0.001.
[2]Two-tailed t-test with unequal variances and [+]α=0.05.

4.3. Sense-of-Presence

The dependent variable was the sense-of-presence score on the Steed-Usoh-Slater Presence Questionnaire. We added two questions on the participants' perception of their self-avatars.

- How much did you associate with the visual representation of yourself (your avatar)? During the experience, I associated with my avatar (1. not very much… 7. very much)

- How realistic (visually, kinesthetically, interactivity) was the visual representation of yourself (your avatar)? During the experience, I thought the avatar was (1. not very realistic… 7. very realistic)

Mean SUS Sense of Presence Score
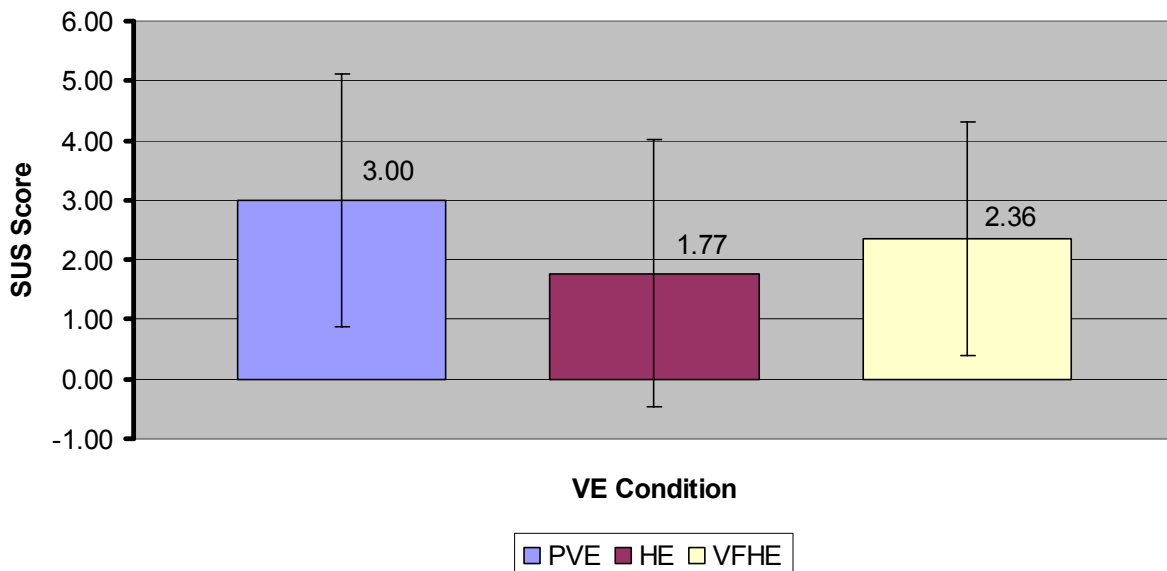
Table 4 – Steed-Usoh-Slater sense-of-presence scores

	Sense-of-presence score (0..6)	
	Mean	S.D
Purely Virtual Environment	3.00	2.12
Hybrid Virtual Environment	1.77	2.24
Visually Faithful Hybrid Environment	2.36	1.95

Table 5 – Self-avatar questions scores

	Self-avatar association (1..7)		Self-avatar realism (1..7)	
	Mean	S.D.	Mean	S.D.
Purely Virtual Environment	4.43	1.60	3.64	1.55
Hybrid Environment	4.79	1.37	4.57	1.78
Visually Faithful Hybrid Environment	4.64	1.65	4.50	1.74

Table 6 – Sense-of-presence between groups

	Between groups sense-of-presence	
	t-test	p-value
PVE vs. VFHE	1.29[1]	0.21
VFHE vs. HE	1.21[1]	0.24
PVE vs. HE	1.44[2]	0.16

[1]One-tailed t-test with unequal variances.
[2]Two-tailed t-test with unequal variances.

4.4. Other Factors

There was no significant difference in simulator sickness and spatial ability between groups.

Spatial ability and task performance were negatively correlated ($r = -0.31$ [small patterns], $r = -0.38$ [large patterns]).

Table 7 – Simulator sickness and spatial ability between groups

	Simulator Sickness		Spatial Ability	
	t-test[1]	p-value	t-test[1]	p-value
PVE vs. VFHE	1.16	0.26	-1.58	0.13
PVE vs. HE	0.49	0.63	-1.41	0.17
VFHE vs. HE	-0.57	0.58	0.24	0.82

[1]Two-tailed t-test with unequal variances.

5. Discussion

5.1. Task Performance

For small and large patterns, both VFHE and HE task performances were significantly better than PVE task performance (Table 1). The difference in task performance between the HE and VFHE was not significant at the $\alpha=0.05$ level (Table 3).

As expected, performing the block-pattern task took longer in any VE than it did in the RSE. The PVE participants took about **three times** as long as they did in the RSE. The HE and VFHE participants took about twice as long as they did in the RSE.

We accept the task performance hypothesis; interacting with real objects significantly affected task performance over interacting with virtual objects.

In the SUS Presence Questionnaire, participants were asked how well they thought they achieved the task, from 1 (not very well) to 7 (very well). The VFHE participants responded significantly ($t_{27}=2.23$, p=0.0345) higher (M=5.43, s.d.=1.09) than PVE participants (M=4.57, s.d.=0.94).

For the case we investigated, *interacting with real objects provided a quite substantial performance improvement over interacting with virtual objects for cognitive manual tasks*. Although task performance in the VEs was substantially worse than in the RSE, the task performance of HE and VFHE participants was significantly better than for PVE participants.

There is a slight difference between HE and VFHE performance (Table 3, p=0.055) for large patterns, but overall, avatar visual fidelity did not affect task performance. The significantly poorer task performance when interacting with virtual objects leads us to believe that the same hindrances would affect task learning, training, and practice.

5.2. Sense of Presence

Although interviews showed visually faithful self-avatars (VFHE) were preferred, there was no statistically significant difference in sense-of-presence compared to those presented a generic self-avatar (HE and PVE). There were no statistically significant differences at the α=0.05 level between any of the conditions for all sense-of-presence questions.

We reject the sense-of-presence hypothesis; a visually faithful self-avatar did not increase sense-of-presence in a VE, compared to a generic self-avatar. The presence of real objects did not increase sense-of-presence.

Slater cautions against the use of the SUS Questionnaire to compare presence across VE conditions, but also points out that no current questionnaire appears to support such comparisons. Just because we did not see a presence effect does not mean that there was none.

5.3. Participant Response to the Self-Avatar

In the analysis of the post-experience interviews, we identified trends in the participants' responses. When reviewing the results, please note that not every participant had a response to a

question that could be categorized. In fact, most participants spent much of the interview explaining how they felt the environment could be improved, regardless of the question!

The post-experience interviews suggests that many participants with generic avatars (HE and PVE) noted that the "avatar moved when I did" and gave a high mark to the self-avatar questions. In fact, all comments on avatar realism from PVE and HE participants related to motion accuracy.

- "It was pretty normal, it moved the way my hand moved. Everything I did with my hands, it followed."

- "The only thing that really gave me a sense of really being in the virtual room was the fact that the hands moved when mine moved, and if I moved my hand, the room changed to represent that movement."

- "Being able to see my hands moving around helped with the sense of 'being there'."

Some participants with visually faithful avatars (VFHE) said, "Yeah, I saw myself" and gave an equally high mark to the avatar questions. This resulted in similar scores to the questions on self-avatar realism. In fact, all comments on avatar realism from VFHE participants related to visual accuracy.

- "Nice to have skin tones, yes (I did identify with them)"

- "Yeah, those were my hands, and that was cool... I was impressed that I could see my own hands"

- "Appearance looked normal, looked like my own hands, as far as size and focus looked absolutely normal… I could see my own hands, my fingers, the hair on my hands"

From the interviews, we conclude participants who commented on the visual fidelity of their self-avatar assumed that its movement would also be accurate. We believe that visual fidelity encompasses kinetic fidelity.

In hindsight, the different components of the self-avatar (appearance, movement, and interactivity) should perhaps have been divided into separate questions. Steed, one of the designers of the SUS Questionnaire, suggested that the cognitive load of the block task could make it hard to detect the relatively smaller differences in the sense-of-presence measures. Regardless of condition, the responses had a movement first, appearance second trend. We hypothesize *kinematic fidelity of the avatar is significantly more important than visual fidelity for sense-of-presence*. We believe the impact of self-avatar visual fidelity on sense-of-presence might not be too strong.

Perhaps two quotes sum up the visually-faithful self-avatars best:

- "I thought that was really good, I didn't even realize so much that I was virtual. I didn't focus on it quite as much as the blocks. "

- "I forget… just the same as in reality. Yeah, I didn't even notice my hands."

5.4. Debriefing Trends

Task Performance.

- Among the reconstruction system participants (HE and VFHE), 75% noticed the reconstruction errors and 25% noticed the reconstruction lag. Most complained of the

limited field of view of the working environment. Interestingly, the RSE had a similar

limited working volume and field of view, but no participant mentioned it.

- 93% of the PVE and 13% of the HE and VFHE participants complained that the

  interaction with the blocks was unnatural.

- 25% of the HE and VFHE participants felt the interaction was natural.

Sense-of-Presence. Participants in all VE groups commented that the following increased their

VE sense-of-presence:

- Performing the task.

- Seeing a self-avatar.

- Virtual objects in the room (such as the painting, plant, and lamp), even though they had

  no direct interaction with these objects.

When asked what factors *increased* their VE sense-of-presence:

- 26% of HE and VFHE participants said having the real objects and tactile feedback.

- 65% of VFHE and 30% of HE participants said that their self-avatar "looked real".

When asked what factors *decreased* their VE sense-of-presence:

- 43% of PVE participants said the blocks not being there or behaving as expected.

- 11% of HE and VFHE participants said manipulating real objects because "they reminded

  them of the real world."

- 75% of HE and VFHE participants said the reconstruction errors, lag, and field of view.

Finally, VFHE participants reported feeling comfortable with the task significantly more quickly than PVE participants (T_{26} = 2.83, p=0.0044) at the α=0.01 level. Participants were comfortable with the workings of the VE almost an entire practice pattern earlier (1.50 to 2.36 patterns).

Overall. The following trends were consistent with previous research or our VE experiences:

- Working on a task heightened sense-of-presence.

- Interacting with real objects heightened sense-of-presence.

- VE latency decreased sense-of-presence.

5.5. Observations

The interaction to rotate the block was the primary component in the difference in times between VE conditions. The typical problem solving method was to pick up a block, rotate it, and check if the new face matched the desired pattern. If it did not match, rotate again. If it matched, place the block in the appropriate place and get the next block. The secondary component of task performance was the selection and placement of a block.

These factors were improved through the tactile feedback, natural interaction, and motion constraints of handling real blocks.

Using the one-size-fits-all pinch gloves had some unexpected fitting and hygiene consequences, even in the relatively small fourteen-participant PVE group.

- Two participants had large hands. They had difficulty fitting into the gloves.

- Two participants had small hands. They had difficulty registering pinching actions because the gloves' sensors were not positioned appropriately.

- One participant became nauseated and quit midway through the experiment. The pinch gloves became moist with sweat, and were a hygiene issue for subsequent participants.

We also saw evidence that the misregistration between the real and virtual actions in the PVE affected participant's actions. Recall that while the participant made a pinching gesture to pick up a block, visually they saw the avatar hand grasp a virtual block (Figure 10). This misregistration caused 25% of the participants to forget the pinching gesture and try a grasping action (which at times did not register with the pinch gloves). If the experimenter observed this behavior, he reminded the participant to make pinching motions to grasp a block.
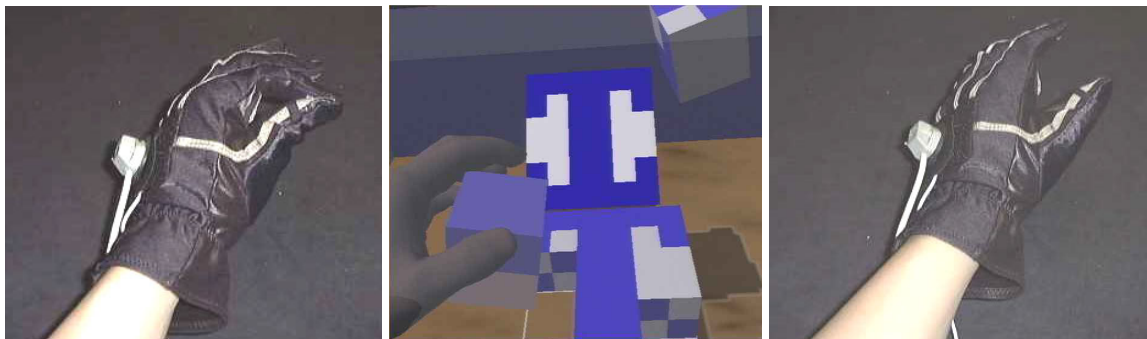
The PVE embodied several interaction shortcuts. For example, blocks would float in midair if the participant released the block more than six inches above the table. This eased block rotation and allowed a select, rotate, release mechanism similar to a ratchet wrench. Some participants, in an effort to maximize efficiency, learned to grab blocks and place them all in midair before the beginning of a pattern. This allowed easy and quick access to blocks. The included shortcuts were carefully considered to assist in interaction, yet led to adaptation and learned behavior.

Most participants mentally subdivided the target pattern and worked on matching one subsection at a time. Each block was picked up and rotated until the desired face was found. Some participants noted that this rotation could be done so quickly in the RSE that they just randomly spun each block to find a desired face. In contrast, two PVE and one HE participant remarked that the slower interaction of block rotation in the VE influenced them to memorize the relative orientation of the block faces to improve performance. For training applications, participants developing behaviors inconsistent with their real world approach to the task could be detrimental to effectiveness or even dangerous.

Manipulating real objects also benefited from natural motion constraints. Tasks such as placing the center block in a nine-block pattern and closing gaps between blocks were easily done with real objects. In the PVE condition (all virtual objects), these interaction tasks would have been difficult and time-consuming. We provided snapping upon release of a block to alleviate these handicaps, but the inclusion of artificial aides might be questionable if the goal of the system was learning or training.

6. Conclusions

Interacting with real objects significantly improves task performance over interacting with virtual objects in spatial cognitive tasks, and more importantly**,** it brings performance measures closer to that of doing the task in real space. *Handling real objects makes task performance and interaction in the VE more like the actual task.*

Further, the way participants perform the task in the VE using real objects is more similar to how they would do it in a real environment. The motion constraints and tactile feedback of the real objects provide additional stimuli that create an experience much closer to the actual task than one with purely virtual objects. Even in our simple task, we saw evidence that manipulating virtual objects sometimes caused participants to incorrectly associate interaction mechanics and develop VE-specific approaches. Training and simulation VEs try to recreate real experiences and would benefit from having the participant manipulate as many real objects as possible.

Motion fidelity is more important than visual fidelity for self-avatar believability. We believe that a visually faithful self-avatar is better than a generic self-avatar, but from a sense-of-presence standpoint, the advantages do not seem very strong. We suggest designers focus their efforts on tracking and animation rather than on rendering quality for immersive VEs. Texture mapping the self-avatar model with captured images of the user would be a big (and relatively straight forward) step towards increased visual fidelity and immersion.

7. Future Work

Does interacting with real objects expand the application base of VEs? We know that the purely virtual aspect of current systems has limited the applicability of VR to some tasks. We look to identify the types of tasks that would most benefit from having the user handle real objects.

Which aspects of a self-avatar are important for presence, and specifically does visual fidelity affect presence in VEs? We believe it does. Yet even if this is true, how strong an effect does it have? Even though our study does not show a significant difference in presence, the participant

interviews leads us to believe there is some consequence. Future work would involve identifying tasks and measures that can isolate the effect of self-avatar visual fidelity on presence.

8. Acknowledgements

9. Bibliography

[1]   F. Brooks Jr. "What's Real About Virtual Reality?" *IEEE Computer Graphics and Applications,* Vol 19, No. 6, pp. 16-27, 1999.

[2]   I. Sutherland. "The Ultimate Display", In *Proceedings of IFIP* 65, Vol 2, pp 506, 1965.

[3]   M. Slater and M. Usoh. "The Influence of a Virtual Body on Presence in Immersive Virtual Environments", *VR 93, Virtual Reality International, Proceedings of the Third Annual Conference on Virtual Reality*, London, Meckler, 1993, pp 34-42.

[4]   M. Usoh, K. Arthur, *et al*. "Walking > Virtual Walking> Flying, in Virtual Environments", *Proceedings of SIGGRAPH 99*, pp. 359-364, Computer Graphics Annual Conference Series, 1999.

[5]   J. Mortensen, V. Vinayagamoorthy, M. Slater, A. Steed, B. Lok, and M. Whitton. "Collaboration in Tele-Immersive Environments", In *Proceedings of Eighth Eurographics Workshop on Virtual Environments* (EGVE 2002) on May 30-31, 2002.

[6]   M. Slater and M. Usoh. "Body Centred Interaction in Immersive Virtual Environments", in N. Magnenat Thalmann and D. Thalmann, Eds., *Artificial Life and Virtual Reality*, pp. 125-148, John Wiley and Sons, 1994.

[7]   D. Pertaub, M. Slater, and C. Barker. "An Experiment on Fear of Public Speaking in Virtual Reality", *Medicine Meets Virtual Reality 2001*, pp. 372-378, J. D. Westwood *et al*. (Eds) IOS Press, ISSN 0926-9630.

[8]   A. Hilton, D. Beresford, T. Gentils, R. Smith, W. Sun, and J. Illingworth. "Whole-Body Modeling of People from Multiview Images to Populate Virtual Worlds", *The Visual Computer*, Vol. 16, No. 7, pp. 411-436, 2000. ISSN 0178-2789.

[9]   K. Hinckley, R. Pausch, J. Goble, and N. Kassell. "Passive Real-World Interface Props for Neurosurgical Visualization", In *Proceedings of the 1994 SIG-CHI Conference*, pp 452-458.

[10] D. Bowman and L. Hodges. "An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments", In 1997 ACM Symposium on Interactive 3-D Graphics, pp. 35-38 (April 1997). ACM SIGGRAPH. Edited by Michael Cohen and David Zeltzer. ISBN 0-89791-884-3.

[11] C. Hand. A Survey of 3-D Interaction Techniques, *Computer Graphics Forum*, Vol. 16, No. 5, pp. 269-281 (1997). Blackwell Publishers. ISSN 1067-7055.

[12] R. Lindeman, J. Sibert, and J. Hahn. "Hand-Held Windows: Towards Effective 2D Interaction in Immersive Virtual Environments", In *IEEE Virtual Reality*, 1999.

[13] B. Lok. "Online Model Reconstruction for Interactive Virtual Environments", In *Proceedings 2001 Symposium on Interactive 3-D Graphics*, Chapel Hill, N.C., 18-21, March 2001, pp. 69-72, 248.

[14] D. Wechsler. The Measurement of Adult Intelligence, 1st Ed., Baltimore, MD: Waverly Press, Inc. 1939.

[15] M. Usoh, E. Catena, S. Arman, and M. Slater. "Using Presence Questionnaires in Reality", *Presence: Teleoperators and Virtual Environments*, Vol. 9, No. 5, pp. 497-503.

Figure 1 - Image of the wooden blocks manipulated by the participant to match a target pattern.

Figure 2 – Each participant performed the task in the RSE and then in one of the three VEs.

Figure 3 – Real Space Environment (RSE). Participant watches a small TV and manipulates wooden blocks to match the target pattern.

Figure 4 – Purely Virtual Environment (PVE). Participant wore tracked pinchgloves and manipulated virtual objects.

Figure 5 – Hybrid Environment (HE). Participant manipulated real objects while wearing dishwashing gloves to provide a generic avatar.

Figure 6 – Visually Faithful Hybrid Environment (VFHE). Participants manipulated real objects and were presented with a visually faithful self-avatar.

Figure 7 – VE for all three virtual conditions.

Figure 8 - Mean time to correctly match the target pattern in the different conditions.

Figure 9 - Mean SUS sense-of-presence questionnaire scores for the different VEs.

Figure 10 – The participant pinches (left) to pick up a block (center). Midway through the experiment, some participants started using a grabbing motion (right).