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

Benjamin Lok
University of North Carolina at Charlotte
bclok@cs.uncc.edu

Samir Naik, Mary Whitton,
Frederick P. Brooks Jr.
University of North Carolina at Chapel Hill
[naiks, whitton, brooks]@cs.unc.edu

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 effectiveness for certain tasks? Further, does the visual fidelity of the virtual objects affect performance? If participants spent most of their time and cognitive load on learning and adapting to interacting with a purely virtual system, this could reduce the overall effectiveness of a VE.

We conducted a study that investigated how handling real objects and self-avatar visual fidelity affects performance 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.

1. Introduction

1.1. Motivation

Conducting design evaluation and assembly feasibility evaluation tasks in immersive virtual environments (VEs) enables designers to evaluate and validate multiple alternative designs more quickly and cheaply than if mock-ups are built and more thoroughly than can be done from drawings. Design review has become one of the major productive applications of VEs [1]. Virtual models can be used to study the following important design questions:

- Can an artifact readily be assembled?
- Can repairers readily service it?

The ideal VE system would have the participant fully believe he was actually performing a task. In the assembly verification example, parts and tools would have mass, feel and look 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. Imagine trying to simulate a task as basic as unscrewing an oil filter from an engine in such a VE!

Interacting with purely virtual objects could impose three limiting factors on VEs:

- Limits the types of feedback, such as constraints and haptics, the system can provide to the user.
- The VE representation of real objects (*real-object avatars*) is usually stylized and not necessarily visually faithful to the object itself.
- Hampers real objects (including the user), naturally interacting with virtual objects.

This work investigates the impact of the first two factors on task performance in a spatial cognitive task. These factors might hinder training and performance in tasks that require haptic feedback and natural interaction. As opposed to perceptual motor tasks (e.g., pick up a pen), cognitive tasks require problem-solving decisions on actions (e.g., pick up a *red* pen). Most design verification and training tasks are cognitive.

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 registered 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, one that could handle *dynamic real objects*, would be effective in providing natural interactivity and visually-faithful self-avatars. In turn, this should improve task performance.

The advantages of interacting with real objects could enable applying VEs to tasks that are hampered by using all virtual objects. We believe spatial cognitive manual tasks, common in simulation and training VEs, would benefit from incorporating real objects. These tasks require problem solving through manipulating objects while maintaining mental relationships among them.

2. Previous Work

The user is represented within the VE with a self-avatar, either 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 in the commercial packages. Although these models contain a substantial amount of detail, they do not visually match each specific participant's appearance.

Researchers believe that providing generic self-avatars substantially improves sense-of-presence over providing no self-avatar [3][4]. However, they 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 concludes, "Substantial potential presence gains can be had from tracking all limbs and customizing [self-]avatar appearance [5]." If the self-avatar visual fidelity might affect sense-of-presence, might it also affect task performance?

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. The human body's deformability and numerous degrees of freedom makes presenting an accurate representation of the participant's pose difficult.

Matching the virtual look to the physical reality is difficult to do dynamically, though static-textured, personalized self-avatars are available through commercial systems, such as the AvatarMe system [5].

Ideally, a participant would interact with the VE in the same way as he would in a real world situation. The VE system would understand and react to expressions, gestures, and motion. The difficulty is in capturing this information for rendering and simulation input.

The fundamental interaction problem is that most things are not real in a virtual environment. 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 device for a specific type of interaction, such as tracking a toy spider registered with a virtual spider [7]. This typically improves interaction affordance, so that the participant interacts with the system in a more natural manner. For example, augmenting a doll's head with sliding rods and trackers enables doctors to more naturally select cutting planes for visualizing MRI data [8]. 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 [9], interaction methodologies [10], and 3-D GUI widgets and physical interaction [11].

3. User Study

3.1. Study Goals

This was part of a larger study that examined the effects of incorporating real objects into VEs. We started off trying to study the following: For cognitive tasks,

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

To test this, we employed a hybrid system that can incorporate dynamic real objects into a VE. It uses multiple cameras to generate virtual representations of real objects at interactive rates [12]. This allowed us to investigate how performance on cognitive tasks, i.e. time to complete, is affected by interacting with real versus virtual objects. The results will be useful for training and assembly verification, as they often require the user to solve problems 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. Thus the system also allows investigating whether a visually faithful self-avatar, as opposed to a generic self-avatar, increases task performance. The results will provide insight into the need to invest the additional effort to use high visual fidelity self-avatars.

3.2. Task Description

We sought to abstract tasks common to VE design applications. Through surveying production VEs [1], we noted that a substantial number involved spatial cognitive manual tasks.

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 participant physical variability.

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 a real-space task 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 [13]. The block-design component measures reasoning, problem solving, and spatial visualization.

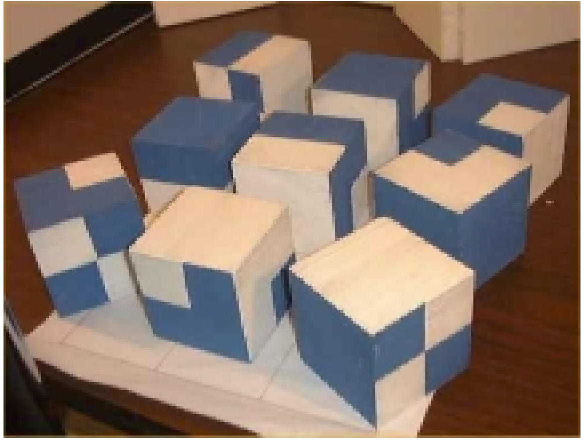

Figure 1 - Image of the wooden blocks manipulated by the participant to match a target pattern.

In the standard WAIS block design task, participants manipulate one-inch cubes to match target patterns. As the WAIS test is copyrighted, we modified the task to still require cognitive and problem solving skills while focusing on interaction methodologies. Also, the small one-inch cubes of the WAIS would be difficult to manipulate with purely virtual approaches and hamper the conditions that used the reconstruction system due to reconstruction error. We increased the size of the blocks to three-inch cubes, as shown in Figure 1.

Participants manipulated four or nine of identical wooden blocks to make the top face of the blocks match a target pattern. Each cube had its faces painted with the six patterns that represented the possible quadrant-divided white-blue patterns.

There were two sizes of target patterns, *small* four-block patterns in a two-by-two arrangement, and *large* nine-block patterns in a three-by-three 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 in a VE condition. 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 with a generic self-avatar (purely virtual environment - PVE)
- Real objects with a generic self-avatar (hybrid environment - HE)
- Real objects with a visually faithful self-avatar (visually-faithful hybrid environment – VFHE)

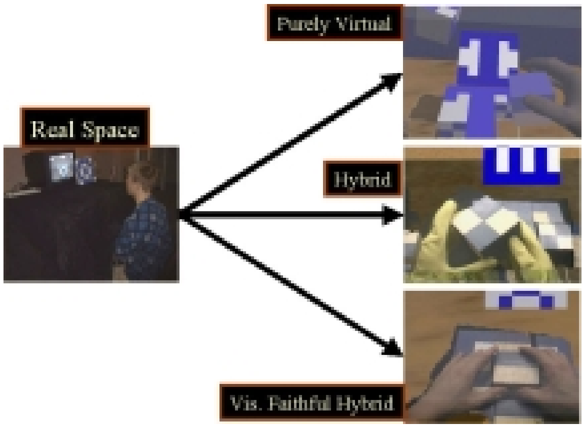

Figure 2 – Each participant performed the task in the RSE and then in one of the three VEs.

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 interaction.

The participants were randomly assigned to one of the three groups, 1) RSE then PVE, 2) RSE then HE, or 3) RSE then VFHE (Figure 2).

Real Space Environment (RSE). The participant sat at a desk (Figure 3) with nine wooden blocks inside a rectangular enclosure. The side facing the participant was open and the whole enclosure was draped with a dark cloth. Two small lights lit the inside of the enclosure.

A television placed atop the enclosure displayed the video feed from a “lipstick 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.

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

Purely Virtual Environment (PVE). Participants stood at a four-foot high table, and wore Fakespace Pinchgloves, each tracked with Polhemus Fastrak trackers, and a Virtual Research V8 head-mounted display (HMD) (Figure 4).

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

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 (its color was a neutral gray).

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

Releasing the block within six inches of the workspace surface caused the block to snap into an unoccupied position in a three by three grid on the table. This reduced the fine-grained interaction that would have artificially inflated the time to complete the task. Releasing the block away from the grid caused it to simply 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).

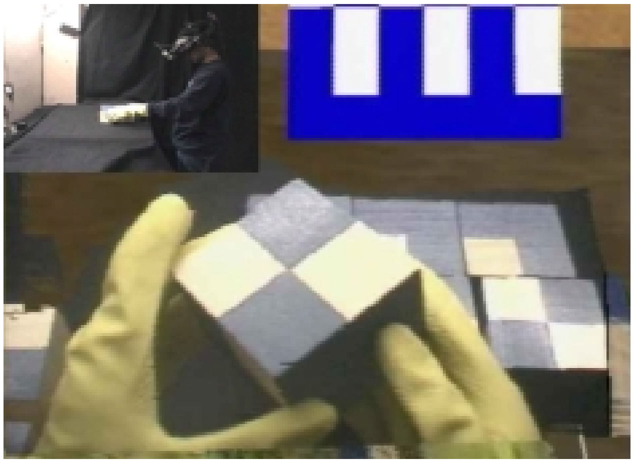

Figure 5 – Hybrid Environment (HE). Participant manipulated real objects while wearing dishwashing gloves to provide a generic avatar.

Visually-Faithful Hybrid Environment (VFHE). Participants wore only the HMD. Otherwise, this condition similar to the HE. The self-avatar was visually faithful, as the shape reconstruction was texture-mapped with images from a HMD mounted camera. The participant saw an image of his own hands (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 participant wore a Virtual Research V8 HMD (640x480 resolution in both eyes) that was tracked with the UNC HiBall tracking system.

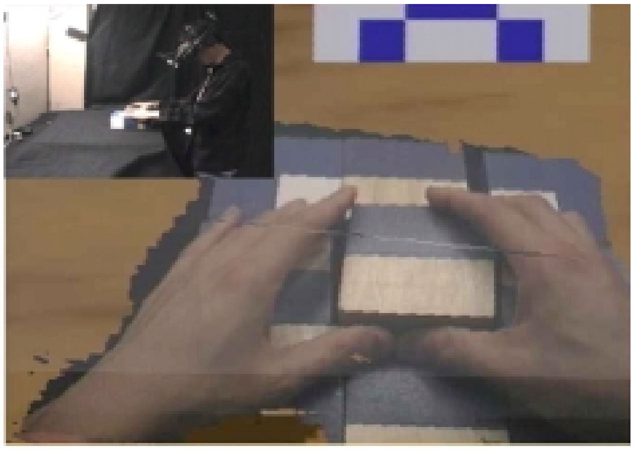

Figure 6 – Visually Faithful Hybrid Environment (VFHE). Participants manipulated real objects and were presented with a visually faithful self-avatar.

The PVE ran on one rendering pipe at twenty FPS. The HE and VFHE ran on four rendering pipes at twenty FPS for virtual objects, and twelve FPS for reconstructing real objects. The reconstruction system used 4 cameras and had an estimated latency of 0.3 seconds and 1 cm reconstruction error.

Figure 7 – VE for all three virtual conditions.

Rationale for Conditions. We expect a participant's RSE (no VE equipment) performance would produce the best results, as the interaction and visually fidelity were optimal. Thus, we compared how closely a participant's task performance in VE was to their RSE task performance.

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, and doing the task in the RSE allowed participants to become proficient without spending additional time in the VE. We limited VE time to fifteen minutes, as many pilot subjects complained of fatigue after that amount of time.

The PVE is a plausible approach to the task with current technology. All the objects were virtual, and interactions were accomplished with specialized equipment and gestures. The difference in task performance between the RSE and the PVE corresponded to the impedance of interacting with virtual objects.

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

The VFHE evaluates the cumulative effect of both real object interaction and visually faithful self-avatars on performance. We were interested in seeing how close participants' performance in our reconstruction system would be to their ideal RSE performance.

3.4. Measures

Task Performance. Participants were timed on replicating *correctly* the target pattern. We also 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. Each participant eventually completed every pattern correctly.

Other Factors. We also measured sense-of-presence, spatial ability and simulator sickness by using the Steed-Usoh-Slater Presence Questionnaire (SUS) [14], Guilford-Zimmerman Aptitude Survey, Part 5: Spatial Orientation and the Kennedy – Lane Simulator Sickness Questionnaire respectively.

Participant Reactions. At the end of the session, we interviewed each participant on their impressions of their experience. Finally, we recorded self-reported 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 entered the room with the real space environment (RSE) setup. The participant was presented with the wooden blocks and was instructed on the task. The participant was also told that they would be timed, and to examine the blocks and become comfortable with moving them. The cloth on the enclosure was lowered, and the TV turned on.

The participant was given a series of six practice patterns, three small (2x2) and then three large (3x3). The participant was told the number of blocks involved in a pattern, and to notify the experimenter when they were done. After the practice patterns were completed, a series of six test patterns were administered, three small and three large. Between patterns, the participant was asked to randomize the blocks' orientations.

Though all participants saw the same twenty patterns, the order of the patterns that each participant saw was unique (six real space practice, six real space timed, four VE practice, four VE timed).

We recorded the time required to complete each test pattern correctly. If the participant misjudged the completion of the pattern, we noted this as an error and told the participant that the pattern was not yet complete, and to continue working on the pattern. We did not stop the clock on errors. The final time was used as the task performance measure for that pattern.

Virtual Space. Next, the participant entered a different room where the experimenter helped the participant put on the HMD and any additional equipment particular to the VE condition (PVE – tracked pinch gloves, HE – dishwashing gloves). The participants were randomly assigned to the various conditions.

Following a period of adaptation to the VE, 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 so desired. Only one person (a PVE participant) asked for a break.

Post Experience. Finally, the participant was interviewed about their impressions of and reactions to the session. The debriefing session was a semi-structured interview. The specific questions asked were only starting points, and the interviewer could delve more deeply into responses for further clarification or to explore unexpected conversation paths.

The participant filled out the simulator sickness questionnaire again. By comparing their pre- and post-experience scores, we could assess if their level of simulator sickness had changed while performing the task. Finally, an expanded Slater – Usoh – Steed Virtual Presence Questionnaire was given to measure the participant's sense of presence in the VE.

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 never *noticed* any problems and continued working. We noted long or repeated tracking failures, and participants who were tall (which gave the head tracker problems) were allowed to sit to perform the task. None of the tracking failures appeared to significantly affect the task performance time.

On hand were additional patterns for replacement of voided trials, such as if a participant dropped a block onto the floor. This happened twice and was noted.

3.6. Hypotheses

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). *Handling real objects improves task performance.*

Participants represented in the VE by a visually faithful self-avatar (VFHE) will complete the spatial cognitive manual task significantly closer to their RSE task performance than will participants who are represented by a generic self-avatar (PVE, HE). *Self-avatar visual fidelity improves task performance.*

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). They were primarily male (thirty-three) undergraduate students enrolled at UNC-CH (thirty-one). Participants were 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 (M=2.85, s.d.=1.26), on [1..7] scales. There were no significant differences between the groups.

During the recruiting process, we required participants to have taken or be currently enrolled in a higher-level mathematics course (equivalent of a Calculus 1 course). This greatly reduced participant spatial ability variability, and in turn reduced task performance variability.

4.2. Experiment Data

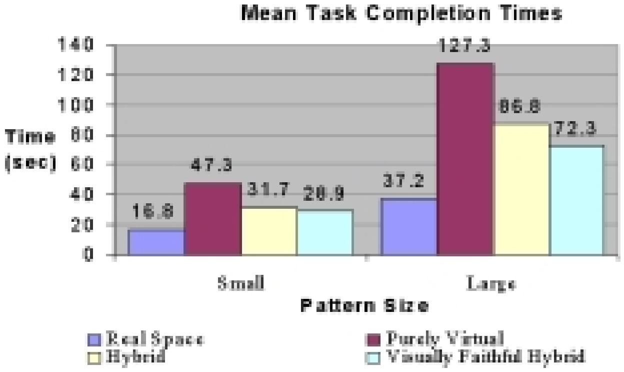

Figure 8 - Mean time to correctly match the target pattern in the different conditions.

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. We use a two-tailed t-test with unequal variances and an $\alpha=0.05$ level for significance unless otherwise stated.

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 – Between groups task performance

	Small Pattern		Large Pattern	
	t-test	p-value	t-test	p-value
PVE-RSE vs. VFHE-RSE	3.32	0.0026^{**}	4.39	0.00016^{***}
PVE-RSE vs. HE-RSE	2.81	0.0094^{**}	2.45	0.021^{*}
VFHE-RSE vs. HE-RSE	1.02	0.32	2.01	0.055^{+}

Significant at the $^{*}\alpha=0.05$, $^{**}\alpha=0.01$, $^{***}\alpha=0.001$

$^{+}$ - requires further investigation

4.4. Other Factors

Sense-of-presence, simulator sickness, and spatial ability were not significantly different between groups. A full analysis of the sense-of-presence results will be discussed in subsequent publications. Spatial ability was moderately correlated (r = -0.31 for small patterns, and r = -0.38 for large patterns) with performance.

Table 3 – Between groups sense-of-presence, simulator sickness, and spatial ability

		PVE vs. VFHE	PVE vs. HE	HE vs. VFHE
Sense-of-Presence	t-test	1.10	1.64	0.64
	p-value	0.28	0.11	0.53
Simulator Sickness	t-test	1.16	0.49	-0.57
	p-value	0.26	0.63	0.58
Spatial Ability	t-test	-1.58	-1.41	0.24
	p-value	0.13	0.17	0.82

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 2).

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, on a scale from 1 (not very well) to 7 (very well). The VFHE participants responded significantly higher (M=5.43, s.d.=1.09) than PVE (M=4.57, s.d.=0.94) participants (t_{27}=2.23, p=0.0345).

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 all the VE conditions 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 2, p=0.055), and we do not have a hypothesis as to the cause of this result. This is a candidate for further investigation.

The significantly poorer task performance when interacting with virtual objects leads us to believe that the same hindrances would affect learning, training, and practicing the task.

Handling real objects makes task performance and interaction in the VE more like the actual task.

Although interviews showed visually faithful self-avatars (VFHE) were preferred, *there was no statistically significant difference in task performance* compared to those presented a generic self-avatar (HE and PVE).

We reject the self-avatar visual-fidelity hypothesis; a visually faithful self-avatar did not improve task performance in a VE, compared to a generic self-avatar.

5.3. Debriefing Trends

- Among the reconstruction system participants (HE and VFHE), 75% noticed the reconstruction errors and 25% noticed the lag. Most complained of the limited field of view of the working environment. Interestingly, the RSE had a similar 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.

- 65% of VFHE and 30% of HE participants commented that their self-avatar “looked real”.
- 43% of PVE participants commented on the blocks not being there or behaving as expected.

Finally, participants were asked how many patterns they needed to practice on before they felt comfortable interacting with the virtual environment. 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).

5.5. Observations

The interactions to rotate the block dominated the difference in times between VE conditions. The typical methodology was to pick up a block, rotate it, and check if the new face is the desired pattern. If not, rotate again. If it is, place the block and get the next block. The second most significant component of task performance was the selection and placement of the blocks. 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 in the fourteen-participant PVE group.

- Two members had large hands and had difficulty fitting into the gloves.
- Two of the participants had small hands and had difficulty registering pinching actions because the gloves’ sensors were not positioned appropriately.
- One participant became nauseated and quit mid-way through the experiment. The pinch gloves became moist with sweat, and became a hygiene issue for subsequent participants.

Figure 9 – The participant pinches (left) to pick up a block (center). Midway through the experiment, some participants started using a grabbing motion (right).

We also saw evidence that the misregistration between the real and virtual space 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 9). This misregistration caused 25% of the participants to forget the pinching mnemonic 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 common tasks. For example, blocks would float in midair if the participant released the block more than six inches above the table. This eased the rotation of the block 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 inclusion of the shortcuts was carefully considered to assist in interaction, yet led to adaptation and learned behavior.

In the RSE, participants worked on matching the mentally subdivided target pattern 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 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 VE-specific 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 into position 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 this involved adding artificial aides that might be questionable based if the system was used for learning or training a task.

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. Even in our simple task, we saw evidence that manipulating virtual objects sometimes caused participants incorrectly associate interaction mechanics and develop VE-specific approaches,.

Training and simulation VEs are trying to recreate real experiences and would benefit from having the participant manipulate as many real objects as possible. 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 if an object reconstruction system is not used, we believe that instrumenting, modeling and tracking the real objects that the participant will handle would significantly enhance spatial cognitive tasks.

Self-avatar visual fidelity is clearly secondary to interacting with real objects, and probably has little, if any, affect on cognitive task performance. We believe that a visually faithful self-avatar is better than a generic self-avatar, but from a task performance standpoint, the advantages do not seem very strong.

7. Future Work

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

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.

[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. 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.

[5] 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.

[6] 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.

[7] H. Hoffman, A. Carlin, and S. Weghorst. “Virtual Reality and Tactile Augmentation in the Treatment of Spider Phobia”, *Medicine Meets Virtual Reality 5*, San Deigo, California, January, 1997.

[8] 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.

[9] D. Bowman and L. Hodges. “An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments”, M. Cohen and D. Zeltzer, Eds., In *Proceedings 1997 ACM Symposium on Interactive 3-D Graphics*, April 1997, pp. 35-38.

[10] 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.

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

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

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

[14] 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.