The Effects of Presentation Method and Simulation Fidelity on Psychomotor Education in a Bimanual Metrology Training Simulation

Jeffrey Bertrand* Ayush Bhargava† Kapil Chalil Madathil‡ Anand Gramopadhye§ Sabarish V. Babu¶

Clemson University

Figure 1: Left: Bimanual interaction in a virtual precision metrology training simulation as seen in a head-mounted display (HMD) presentation condition. Right: Viewing and spatially interacting with objects in a stereoscopic large-screen immersive display (LSID) presentation condition.

Abstract

In this study, we empirically evaluated the effects of presentation method and simulation fidelity on task performance and psychomotor skills acquisition in an immersive bimanual simulation towards precision metrology education. In a 2×2 experiment design, we investigated a large-screen immersive display (LSID) with a head-mounted display (HMD), and the presence versus absence of gravity. Advantages of the HMD include interacting with the simulation in a more natural manner as compared to using a large-screen immersive display due to the similarities between the interactions afforded in the virtual compared to the real-world task. Suspending the laws of physics may have an effect on usability and in turn could affect learning outcomes. Our dependent variables consisted of a pre and post cognition questionnaire, quantitative performance measures, perceived workload and system usefulness, and a psychomotor assessment to measure to what extent transfer of learning took place from the virtual to the real world. Results indicate that the HMD condition was preferable to the immersive display in several metrics while the no-gravity condition resulted in users adopting strategies that were not advantageous for task performance.

Index Terms: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities;

*e-mail: jbertra@clemson.edu
†e-mail: ayushb@clemson.edu
‡e-mail:kmadath@clemson.edu
§e-mail:agramop@clemson.edu
¶e-mail: sbabu@clemson.edu

2017 IEEE Symposium on 3D User Interfaces (3DUI)
March 18-19, 2017, Los Angeles, CA, USA
978-1-5090-6716-9/17/$31.00 © 2017 IEEE

1 Introduction

The use of commercial-off-the-shelf tracking and display technology for virtual training applications has become more widespread in recent years. Advancements in computing power and development tools have led to improved fidelity of virtual reality simulations for training and education [24]. However, it is unclear to what extent simulation fidelity and presentation method improves or hinders the task performance and pedagogical benefits of simulation based training. As defined in McMahan's Framework for Interaction Fidelity Analysis [25], *interaction fidelity* is defined as the degree of exactness with which real-world interactions are reproduced in an interactive simulation and *simulation fidelity* refers to the degree of exactness with which real-world physics are reproduced in the simulation. Empirical evaluations of the components of simulation and interaction fidelity are essential for determining the most effective interaction metaphors for virtual training simulations.

Designing 3D interaction techniques affords more opportunities and increased challenges as compared to traditional 2D mouse-based techniques [7]. The benefits include allowing the user to perform natural gestures and actions similar to how they would perform the same actions in the real world. However, these actions are often physically demanding due to performing actions in mid-air with visual and auditory feedback only, and without haptic feedback [23, 2]. Simulations employing haptic feedback may mitigate some of these issues [5]; though haptic devices generally remain prohibitively expensive. A key challenge for researchers is finding the right balance between usability, performance, and effort that the user exerts to accomplish a motor task in virtual reality simulations.

Head mounted display technology is also becoming more widespread and requires systematic evaluation to determine the benefits in real-world training tasks. HMDs enable the user to naturally interact with the virtual environment similar to how they would in a real-world setting. The displays have also been shown to exhibit similar distance underestimation compared to large screen immersive displays [12]. However, users will generally have more experience with interacting in 3D simulations using a monitor or large screen display. Questions remain as to whether users will be able to adapt to the HMD and if it will affect usability and performance.

Suspending the laws of physics in a VR simulation could potentially lead to improved usability which in turn could result in improved learning outcomes, but at the cost of reducing simulation fidelity and lack of symmetry between the virtual reality simulation and its real world analogue. Simplifying the physics properties of simulations in order to promote ease of use, acclimation, and learning in simulations involving complex manual spatial interaction metaphors could negatively affect motor skill acquisition and transfer of learning from the virtual task to real world practice. In order to examine these effects, we conducted an empirical evaluation in a 2×2 experiment design to study the effects of presentation method (large-screen stereoscopic viewing vs. head-mounted display viewing), and simulation fidelity (presence or absence of gravity) on task performance and psychomotor skills learning. We developed a complex virtual reality simulation employing bimanual interaction for teaching and training cognitive and motor aspects of precision measurement to aviation and automotive technical students, which was used as the experiment platform for this empirical evaluation.

Our virtual training simulation teaches technical college students in partnering institutions how to use precision metrology instruments such as calipers and micrometers (see Figure 1). Metrology is the science of measurement and is a fundamental skill that automotive and aviation technicians must have before joining the workforce. The simulation procedure follows a scaffolded learning approach where users are given a large amount of instruction in the beginning and guidance is slowly reduced as the user becomes more proficient in the skills. Taking precise measurements is an inherently two-handed or bimanual task, thus a bimanual interaction metaphor was implemented. There is both a cognitive element in the ability to read and interpret the Vernier scale of the instrument and a motor component in the ability to manipulate the measurement instrument into the best position for taking a precise measurement. Simulations addressing the cognitive element of reading and interpreting the Vernier scale on these instruments have been investigated by Al-Zahrani et al. [1], however few have examined the motor element of using calipers and micrometers. In a manufacturing environment such as an assembly line, quality assurance technicians must perform accurate measurements in rapid succession of various parts thus requiring significant motor training. Students are typically taught these skills in a laboratory environment with the guidance of an instructor. However, lab time is limited and would be better utilized by focusing on more advanced topics. Therefore, if a student came to class already possessing the skill, instructors could dedicate more lab resources to higher order skills. Our virtual metrology simulation was designed to address these issues, but questions remain as to what extent immersive presentation methods and simulation fidelity affects learning outcomes and user performance.

2 Related Work

In order to rigorously evaluate different characteristics of interaction fidelity, McMahan developed the Framework for Interaction Fidelity Analysis (FIFA) to systematically classify aspects of virtual interaction and how they relate to the real-world analogue [25]. The FIFA framework was designed to address the need for researchers to systematically address each aspect of interaction fidelity independently to obtain results that can be analysed accurately. Our previous work examined the effects of dimensional symmetry or the degrees-of-freedom afforded in the simulation on task performance and learning outcomes [4]. Results indicated that higher dimensional symmetry could improve motor performance. Our current study aims to build on our previous work by examining presentation method and simulation fidelity on task performance, and psychomotor skills education, in the context of the virtual metrology training simulation.

Nabiyouni et al. used FIFA to categorize different aspects of their locomotion interfaces, namely the Virtusphere, a large, omnidirectional walking mechanism, gamepad, and real walking [28]. The authors provide a detailed discussion of the effects of interaction fidelity in the context of FIFA and show that moderate fidelity techniques (in this case, the Virtusphere), may perform worse than high fidelity and well-designed low fidelity techniques. McMahan notes that previous research generally does not evaluate both interaction and display fidelity in a controlled manner that avoids confounding the two [25]. Few studies address tasks beyond simple view control tasks; hence there is a need for highly controlled investigations on more complex tasks that users may encounter in an actual virtual reality training simulation.

2.1 Bimanual Interaction

Almost every physical task in which humans partake involves two hands either directly or indirectly. For instance, handwriting was thought to be a unimanual task, but Guiard showed that writing was 20% faster when the participants were able to use their non-dominant hand to manipulate the page [14]. Researchers have suggested that using two hands in parallel could improve the efficiency of task performance rather than performing steps sequentially [8]. Hinckley et al. found that two-handed inputs not only improve efficiency, but they also make spatial input comprehensible to the user since it grounds the user in the interaction space, leading to less disorientation [16]. The researchers also note that novice users were able to operate the two-handed interface effectively within minutes [17], reinforcing previous work by Buxton [8] that users can transfer everyday skills for manipulating tools with two hands to the operation of a computer, with little or no training.

Guiard proposed the Kinematic Chain (KC) Model [14] to systematically classify motor actions using two hands. The basic principles include the non-dominant hand setting a dynamic frame of reference for the dominant hand to make fine motor adjustments [18]. An example is threading a needle. If one is holding the needle in the non-dominant hand, the dynamic frame of reference aids the dominant hand in making fine motor adjustments. Both hands are performing asymmetric actions where gross movements from the non-dominant hand inform the fine movements of the dominant hand. Another principle states that the non-dominant hand initiates the action. VR researchers in bimanual interaction regularly cite Guiard's theoretical work to aid in the design of their simulations [3, 19].

Previous research has demonstrated some of the benefits of using a bimanual interaction metaphor. Leganchuk et al. found that bimanual manipulation could result in both cognitive and motor skills benefits due to the ability to manipulate multiple degrees of freedom simultaneously thus decreasing cognitive load over a unimanual method [20]. The researchers found that as their bimanual task became more cognitively demanding, performance over the unimanual method became more pronounced. Gribnau et al. compared one vs. two-handed input for an abstract assembly task and found that after some amount of training, the two-handed condition performed significantly faster than the single-handed metaphor [13]. Likewise, Schultheis et al. found that their two-handed input metaphor outperformed both mouse based and unimanual wand-based interaction metaphors for a 3D construction task [34]. Balakrishnan et al. showed that as long as the task remains asymmetric, temporal benefits can be demonstrated [3], which reinforces Guiard's KC model. Murayma et al. implemented the SPIDAR-G&G bimanual interaction technique and demonstrated shorter completion times over the unimanual 6-DOF metaphor [27]. Based on the previous work, bimanual methods generally outperform unimanual methods in near-field object manipulation tasks requiring high amounts of precision.

2.2 Psychomotor Skills Acquisition

A key advantage of using immersive virtual reality over mouse and keyboard-based VR is the ability to train motor skills. Devices that track body movements can be employed to train manual dexterous movements. Psychomotor skills involve complex interactions between the perceptual (visual, auditory, tactile) and the motor aspects (proprioceptive and kinesthetic) of manual work [9]. Initial demonstration, practice and feedback are important aspects of teaching psychomotor skills [29]. Lehmann et al. demonstrated that psychomotor skills acquired in their virtual endoscopic surgery trainer (VEST) system were transferrable to physical reality [21]. Parmar et al. compared Desktop VR (DVR) to an HMD configuration and evaluated the cognitive and psychomotor learning outcomes in the context of an electrical circuitry simulation [31, 30]. Results indicated that cognitive benefits were more pronounced in the HMD condition while psychomotor skills improved in the DVR presentation metaphor over the HMD.

2.3 Object Clutching and Gravity

Virtual interaction spaces can be categorized into the motor, control, and visual spaces [2]. The motor space refers to the physical volume where users can reach to interact with virtual objects. The control space is the virtual volume where users can control objects. The visual space is defined as the visual representation of the virtual environment that the user perceives. Virtual hand techniques often couple the visual and motor spaces (sometimes referred to as a 1:1 Control/Display ratio [2]) allowing for a natural interaction metaphor, while mouse-based interaction requires the user to perform a mental rotation and translation between the motor and visual spaces. A clutching (or freezing) mechanism refers to the ability to decouple the motor and control spaces, allowing the user to relocate the control space manually in order to aid object selection. Thus if an object was out of reach within their motor space, a user could initiate a clutching action to move the control space within range of the desired object. However, clutching could increase cognitive load due to an extra step that the user must learn and could lead to fatigue, resulting in decreased user performance [16].

Similar in concept is enabling or disabling gravity. When a user drops an object in free space without the influence of gravity, they would be able to relocate their hands and grasp the object from a more advantageous spot without having to reinitiate a grasping action. Few prior studies have examined the effect of gravity, one example being Slater et al. examining the effects of gravity on the user's sense of presence [35]. However, the authors found no significant differences between the gravity and no-gravity condition on the sense of presence despite the laws of physics being violated. This suggests that users may adapt to non-physical properties of virtual simulations.

2.4 Contributions

This work examines the effects of presentation method with respect to the use of a HMD compared to a LSID on the pedagogical benefits and user performance. It adds to the body of knowledge regarding near-field bimanual motor skills training in immersive virtual training simulations. In addition, this work examines the effect of simulation fidelity with respect to the presence or absence of gravity on user performance and learning outcomes. Furthermore, we aim to greatly add to the existing body of knowledge on the effects of factors pertaining to interaction fidelity in VR, specifically presentation method and simulation fidelity, on presence, satisfaction, and usability.

3 Virtual Metrology Training Simulation Testbed

We have designed a bimanual interaction metaphor for physically interacting with virtual objects affording 6 spatial degrees-of-freedom per hand. Building on our previous work [4], we have enhanced the simulation in order to study the effects of visuo-motor co-location in the context of interaction fidelity as well as the effects of physical fidelity with regards to the presence of gravity on system usability and learning outcomes. The user experienced 3 training procedures and 7 instrument modules including the inside, outside, and depth functions of the calipers as well as the inside, outside, and depth micrometers, and a ruler module used as training. Each module was divided into an Introduction, Guided Practice and Open Exercise phase in a scaffolded learning approach.

3.1 Apparatus

3D object manipulation was facilitated by the Razer Hydra electromagnetic tracking controllers in each condition. In the head-mounted display condition (HMD), the display presentation was an Oculus Rift Development Kit 2 (960×1080 per eye, horizontal FOV = $100°$) and the head tracked, large-screen immersive display (LSID) condition consisted of a stereoscopic (80", 1920×1080, horizontal FOV = $53°$) display. Users were seated approximately 2m from the display. Head-tracking was implemented using the Ascension Flock-of-Birds and stereoscopy was implemented using the Nvidia 3D Vision Pro active stereo glasses (see Figure 1). Head-tracking and stereoscopy have been shown to increase user performance on small-scale spacial inspections [32] and were implemented here so that the HMD condition did not have the advantage of motion-parallax and stereocopy cues on spatial perception over the LSID condition. The Unity game engine was used to develop the simulation and Blender was used to model the 3D objects.

In order for users to manipulate the virtual objects, they were required to press and hold the trigger on the back of the Hydra to select and manipulate the instruments and measurement objects. Holding the button down simulated how the user would have to continuously hold a real instrument. However, if the user determined that they could not rotate their hands far enough to take a proper measurement, they were forced to drop the object and try again. This could result in lower user performance and potentially increase user frustration. In order to solve this issue, we disabled gravity in one condition which let the user release the trigger and would freeze or clutch the virtual object in the air. Then they could relocate their hands and grasp the object from a more advantageous spot. However, Hinckley recommended against implementing a clutching mechanism in some cases because it could lead to user confusion and would require the user to learn and remember an extra step to manipulate objects [16]. Another issue stems from users reverting to a unimanual method of interaction. In our simulation, users could simply leave the measurement object in the air and use their dominant hand to take a measurement. This would result in breaking the kinematic chain that Guiard described for bimanual interaction [14], resulting in users not taking advantage of their non-dominant hand setting a dynamic frame of reference for the dominant hand to perform fine motor adjustments.

3.2 Simulation Procedure

3.2.1 Training

Three training scenarios were designed to teach the user how to interact with the virtual objects via the Razer Hydra motion controllers. The first task was to select a ball and drop it in a basket. The next task was similar to the docking task in work done by Schultheis et al. where users selected a shape such as a cube or a cylinder and had to place the shape in the corresponding notch [34]. Results from their study strongly favored the two-handed interface compared to the unimanual mouse interface for an object docking task. The final task had the user manipulate a torus to the end of a curved rail, emphasizing both location and rotation manipulation skills. Rose et al. developed a similar curved rail task to study training transfer to the real world and found that both real and virtual training resulted in equivalent levels of performance [33]. The first instrument module was the ruler and was considered a training module which was left out of the data analysis.

3.2.2 Introduction Phase

The Introduction phase showed the user what the different parts of each instrument were and how to select and manipulate both the instrument and measurement object at the same time. Text of the instructions were displayed in front of the user on the virtual projector (see Figure 2) and was also read out loud via text-to-speech. Instructions in each condition (HMD, LSID, gravity, no-gravity) were exactly the same to avoid confounds. Additionally, each condition used the same virtual environment with the same object scales.

In order to show the orientation of the user's hands, a 3D model of the Razer Hydra controllers mapped to the exact position of the real controllers. The model was mostly transparent to mitigate occlusion issues. The selector sphere at the end of the controllers indicated where the user could pick up objects. The sphere would highlight when it intersected a selectable object's collision boundaries. While primitives such as cubes and capsules are more performant than mesh-based collision boundaries, this simulation employed polygonal-based collision boundaries that fit the 3D model better than what primitives could approximate.

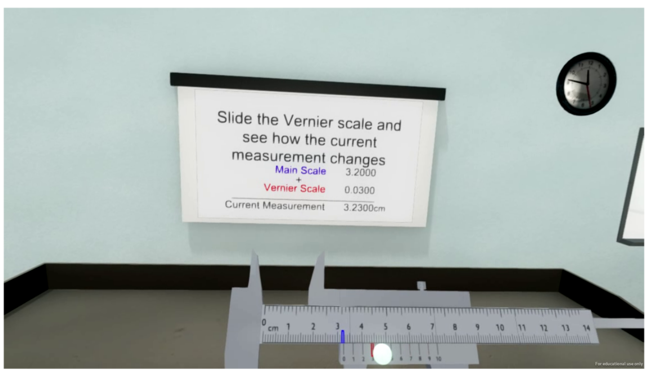

Figure 2: Dynamic instruction shows how to interpret both the main and vernier scales

3.2.3 Guided Practice Phase

The Guided Practice phase led the user step-by-step through the process of taking a precise measurement. The simulation first showed where on the instrument and the measurement object to align either the caliper jaws or the micrometer anvil. It also showed how to open the caliper jaws or micrometer spindle by pressing the thumbstick while the instrument was selected. The next step was to use the lockscrew to lock the measurement in place by pressing one of the face buttons on the Hydra. Finally, interactive highlights showed the user how to read the Vernier scale and were updated dynamically based on where the user had their current measurement set at (see Figure 2). All instructions automatically proceeded to the next step depending on event triggers such as properly calibrating the instrument and using the lockscrew. If the user supplied an answer that was outside the $\pm$1mm tolerance range from the ground truth, the user was asked to try again and the number of attempts were incremented by one.

A method for inputting numbers into the simulation that worked well for both the LSID and HMD conditions was required. Radial menus have been shown to be beneficial in terms of usability and user preference in immersive environments [11]. Therefore, a radial menu was implemented that allowed the user to use the thumbstick on the Hydra to select numbers for their measurements (see Figure 3). Users were shown how to use the radial menu during the Guided Practice phase. The radial menu was attached to the user's hand, taking advantage of what Mine et al. described as *Physical Mnemonics* or storing menus and widgets relative to the user's body [26].

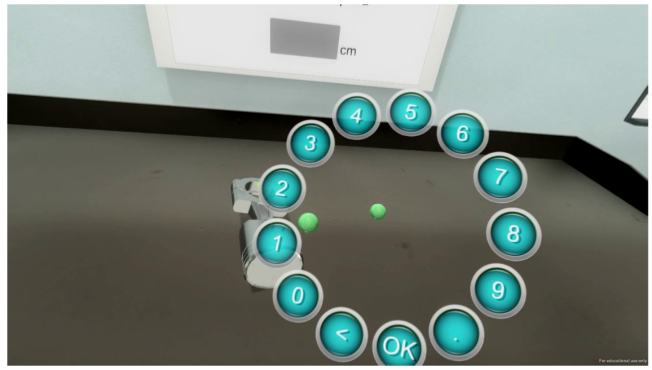

Figure 3: The radial menu used in each condition for inputting measurements

3.2.4 Exercise Phase

Instructions were limited in the Exercise phase and required users to apply what they learned in the Guided Practice phase. They were tasked with measuring real-world objects such as gears and bearings as opposed to abstract objects such as boxes and cylinders in the Guided Practice phase. The simulation provided immediate feedback on whether their answers were correct and were given 5 trials per instrument module. Time to complete was measured by computing the difference between when the instrument appears and when the user submits a correct answer. The objects were randomly scaled after each correct answer. The instructions asked the participants to measure the objects as quickly and accurately as possible.

3.3 After-Action Review Tool

All participant data were saved as XML files at the end of the simulation. The position and orientation of the user's hands and the virtual objects, as well as simulation state were recorded for every frame. We developed an After-Action Review Tool to track the strategies used by participants in each condition. It was also used for data verification and calculating the distances covered by the users' hands and the virtual objects. Graphical elements on the sides of the screen updated in real time as participant data were played back (see Figure 4).

4 Experiment Procedure and Design

In order to examine the effects of presentation method and the presence of gravity on usability and learning outcomes, we have designed a 2×2 study with gravity either enabled or disabled, and an HMD versus a head-tracked, stereoscopic immersive display (LSID). The HMD allows for natural head movements that mimic how one would take measurements in the real world compared to the LSID condition. However, users generally do not have experience with an HMD which could result in worse performance and perceived usability. Disabling gravity could aid in object manipulation and result in improved usability, but this may have an effect on how users accomplish the task and could result in worse performance if the user adopts a unimanual because they will be losing the dynamic frame of reference set by the dominant hand during object manipulation. We observed this behavior in the pilot studies

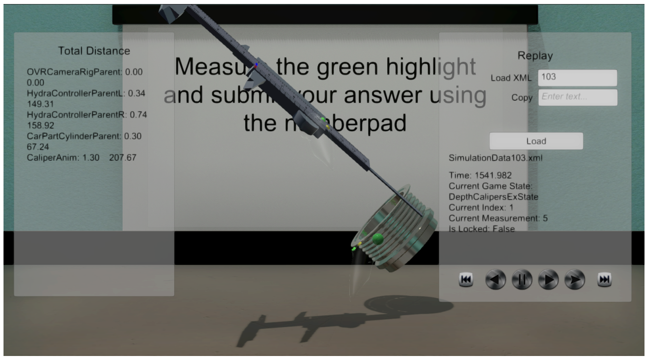

Figure 4: Reviewers could observe strategies that participants employed in the After-Action Review Tool

and we predict participants will also adopt a unimanual approach in the no-gravity condition. These issues led to our research questions and hypotheses.

4.1 Research Questions

1. How does the presentation method in terms of using an HMD or immersive display affect both cognitive and psychomotor learning outcomes?

2. How does presentation method affect user performance and preference?

3. Does enabling or disabling gravity aid the learning process or affect user performance and manipulation strategies?

4.2 Hypotheses

Based on our research questions and previous work, we hypothesize that:

1. Participants will score significantly lower in usability and performance metrics in the LSID condition compared to the HMD condition

2. Participants in the no-gravity condition will score significantly lower in usability and performance metrics

3. Psychomotor learning outcomes will be significantly lower in the no-gravity condition

4.3 Methodology

The experiment procedure is as follows:

1. Participants first read the informed consent and filled out an online demographics survey. The participant was then administered the Cube Comparison mental rotation questionnaire to assess innate mental rotation abilities.

2. The participant took the pre-cognitive assessment questionnaire to determine their prior knowledge of calipers and micrometers.

3. The participant's inter-pupillary distance (IPD) was measured via the Oculus Rift IPD utility.

4. As shown in Figure 1, the user was seated in front of a large-screen display and the IPD and participant ID were entered into the simulation.

5. The participant was then introduced to the controllers and given 3 training tasks in order to become acclimatized to the movement, selection, and manipulation features.

6. Participants interactively experienced 7 modules in the 3D simulation, one for each instrument in this order: Ruler, Outside Calipers (OC), Inside Calipers (IC), Depth Calipers (DC), Outside Micrometers (OM), Inside Micrometers (IM), and Depth Micrometers (DM). Each module was divided into 3 phases consisting of an Introduction, Guided Practice and Exercise phase.

7. Upon completing the simulation, the participant filled out a post-cognitive questionnaire where the questions were similar but not identical to the pre-cognitive assessment questionnaire.

8. The participant then completed the Post-Study System Usability Questionnaire (PSSUQ) [22], Presence Questionnaire [36], and the NASA-TLX [15].

9. Lastly, the participant completed a psychomotor assessment in which they had to use the metrology devices learned in the VR simulation to perform real-world measurements.

4.4 Measures

4.4.1 Pre and Post Cognition Questionnaires

Pre and post cognition questionnaires were administered to participants immediately before and after training in the virtual simulation. Questions were developed by subject matter experts and were categorized into five levels of Bloom's Taxonomy [6] including Knowledge, Comprehension, Application, Synthesis and Evaluation. Questions in the pre-test were similar, but not exactly the same as the post-test. For example, a pre-test Knowledge level question asked the user to identify an inside micrometer and the post-test asked the user to identify a depth micrometer from an image.

4.4.2 Psychomotor Skills Assessment

In order to evaluate skill transference to the real world, a psychomotor assessment was administered at the end of the experiment. Participants were asked to take real-world measurements using the same metrology instruments from the simulation in order to assess training transference to the real world.

4.4.3 Subjective Responses

After interacting with the virtual environment, subjective usability, perceived workload, and presence were measured with the PSSUQ [22], NASA-TLX [15], and the Witmer et al. Presence Questionnaire Ver. 3.0 [36]. The NASA-TLX first asked the user to rate their perception of workload on a 20-point scale in categories including Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration. Next, participants made pair-wise comparisons between each factor. The final weighted score was the sum of the pair-wise counts for each factor multiplied by the normalized scale. Questions on the Presence Survey were rated on a 7-point scale and were categorized into factors including Involvement, Sensory Fidelity, Adaptation/Immersion, and Interface Quality. The PSSUQ usability questions were measured on a 5-point scale and were classified into System Usability, Information Quality, and Interaction Quality.

4.4.4 Virtual Manipulation Performance

Time to complete was recorded for each of the 5 trials in the Exercise phases as well as the Guided Practice phases. Additionally, the number of physical contacts between the virtual instrument and the measurement object were recorded to analyze how many times the objects collided. The simulation did not allow the user to advance to the next step until they supplied an answer that was within a 1mm range from the ground truth. When the user supplied a correct answer, the simulation recorded the user's answer and the reading that the measurement instrument displayed at the time of submission. Otherwise, the number of attempts was incremented and the user was asked to try again. The absolute value of the difference between the reading on the instrument minus the ground truth was defined as the physical difference. It was a function of the user's ability to maneuver the virtual object to the correct position and was a measure of the user's motor skill. A physical difference measure of 0cm would indicate that the user clamped the measurement instrument perfectly. The reading difference was defined as the absolute value of the difference between what the user supplied as an answer minus what the instrument displayed at the time of submission irrespective of the ground truth. The user answer difference was a combination of both the cognitive and motor skill of the user and was measured by taking the absolute value of the ground truth minus the answer supplied by the user.

5 Results

5.1 Participants

Forty-one volunteers (7 female, 34 male) were counted in the analysis. One participant was dropped due to technical issues. Participants were recruited from computing and engineering classes and were monetarily compensated. Ages ranged from 18 to 38 with a mean of 23.1. We asked participants to volunteer only if they had little to no prior usage with calipers and micrometers. The demographics questionnaire asked about prior usage with calipers and micrometers and participants were asked to rate their mastery on a 5-point scale if they had used the instruments before. Seventeen participants (with reported mastery level M = 2.4, SD = 1.3) stated prior use of calipers and most commented that it had been several years since they had used the instrument. Five participants (reported mastery M = 2.8, SD = 1.1) stated prior use of the outside, inside, and depth micrometers. The study lasted approximately 2 hours and the virtual interaction segment lasted approximately 45 minutes.

5.2 Quantitative Analysis

5.2.1 Cognition Scores

A 2 × 2 × 2 mixed model repeated measures ANOVA was performed on the mean scores of each of the levels of Bloom's taxonomy including Knowledge, Comprehension, Application, Synthesis, and Evaluation, with a significance level of α=0.05. The within-subjects effects were the mean pre and post test score results, and the between-subjects effects were the levels of presentation method (LSID vs. HMD) and simulation fidelity (gravity vs. no-gravity) on each dependent variable scores. Interaction effects were also tested between the interaction condition and the pre- and post- test sessions using Bonferroni adjusted alpha method, and in the pairwise comparisons of levels of the between-subjects variables were compared using Tukey HSD method. Mauchly's test of sphericity was used to ensure that error variance in the groups of samples were equivalent. The following sections document the observed effects:

(a) *Knowledge level:*

The mixed model ANOVA revealed a significant main effect of session in that overall, participants in the post-test session (M = 70.83%, SD = 23.85) scored significantly higher with respect to knowledge level questions as compared to the pre-test session (M = 45.23%, SD = 24.37), F(1, 38) = 62.65, $p <$ 0.001, η_p^2 = 0.62.

(b) *Comprehension level:*

The mixed model ANOVA revealed a significant main effect of session in that overall, participants in the post-test session (M = 71.03%, SD = 20.51) scored significantly higher with respect to comprehension level questions as compared to the pre-test session (M = 27.38%, SD = 27.99), F(1, 38) = 74.35, $p <$ 0.001, η_p^2 = 0.66.

(c) *Application level:*

The mixed model ANOVA revealed a significant main effect of session in that overall, participants in the post-test session (M = 66.39%, SD = 26.55) scored significantly higher with respect to application level questions as compared to the pre-test session (M = 19.94%, SD = 21.39), F(1, 38) = 104.248, p $<$ 0.001, η_p^2 = 0.73.

(d) *Synthesis level:*

Here too, the mixed model ANOVA revealed a significant main effect of session in that overall, participants in the post-test session (M = 73.81%, SD = 18.27) scored significantly higher with respect to synthesis level questions as compared to the pre-test session (M = 34.22%, SD = 25.46), F(1, 38) = 94.89, $p <$ 0.001, η_p^2 = 0.71.

(e) *Evaluation level:*

Similar to the other levels, the mixed model ANOVA analysis revealed a significant main effect of session in that overall, participants in the post-test session (M = 94.44%, SD = 12.57) scored significantly higher with respect to evaluation level questions as compared to the pre-test session (M = 34.52%, SD = 32.59), F(1, 38) = 113.215, $p <$ 0.001, η_p^2 = 0.75.

We found that in all the levels of cognition based on Bloom's Taxonomy, overall participants scored significantly higher in the post-test session than in the pre-test session, attesting to the fact that the simulation based training was effective overall in educating the users in precision metrology concepts in all cognitive domains. However, we did not find a significant main effect of simulation fidelity or presentation method or multi-way interactions in our analysis. Figure 5 shows the pre and post cognition scores separated into levels of Bloom's Taxonomy. Error bars in all graphs represent 95% confidence intervals.

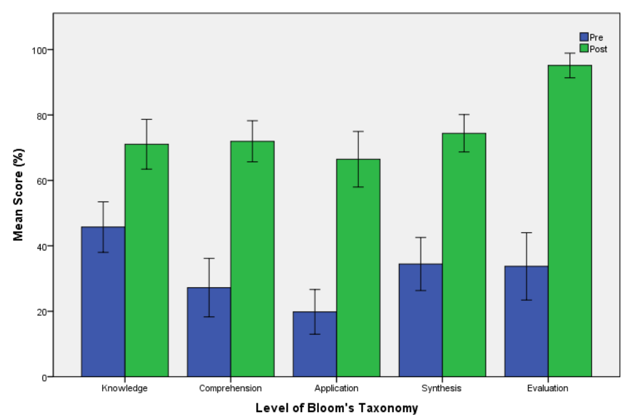

Figure 5: Overall pre and post cognition scores categorized by Bloom's Taxonomy

5.2.2 Psychomotor Skills Assessment

We conducted a 2 × 2 ANOVA on the scores of the participants in the real-world psychomotor assessment test, categorized by Dave's Psychomotor Taxonomy [10] levels of Precision and Articulation. The ANOVA analysis did not reveal a significant main or interaction effect on levels of presentation method or simulation fidelity.

5.2.3 NASA-TLX Workload Assessment Results

We conducted a 2 × 2 Multivariate Analysis of Variance (MANOVA) on the weighted average scores of NASA-TLX Workload Assessment in the dimensions of Mental Demand, Physical Demand, Temporal Demand, Performance and Effort (see Figure 6). The independent factors were presentation method (HMD vs. LSID) and simulation fidelity (gravity vs. no-gravity). The MANOVA analysis revealed a significant main effect of presentation method on the weighted average of mental workload scores, F(1, 41) = 8.68, p = 0.006, η_p^2 = 0.19. Participants in the HMD condition (M = 254, SD = 122.34) rated significantly higher mental workload than participants in the LSID condition (M = 136, SD = 128.30). The MANOVA analysis also revealed a significant main effect of presentation method on the weighted average of perceived performance scores, F(1, 41) = 5.53, p = 0.024, η_p^2 = 0.13. Participants in the HMD condition (M = 292, SD = 99.46) rated significantly higher perceived performance than participants in the LSID condition (M = 195, SD = 158.63). Finally, the MANOVA analysis revealed a significant main effect of presentation method on the weighted average of frustration scores, F(1, 41) = 11.70, p = 0.002, η_p^2 = 0.24. Interestingly participants in the LSID condition (M = 163.81, SD = 130.39) rated significantly higher frustration workload than participants in the HMD condition (M = 43, SD = 84.47), and was the opposite of the results found in the mental and performance workload dimensions.

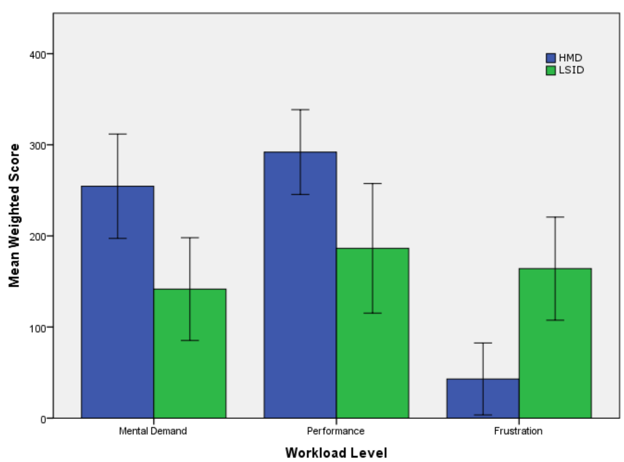

Figure 6: Significant results from the NASA-TLX Workload assessment by presentation method

5.2.4 Presence Results

We conducted a 2 × 2 independent samples ANOVA on the presence dimensions of mean Involvement, Sensory Fidelity, Adaptation/Immersion and Interface Quality scores. The independent factors were presentation method (LSID vs. HMD) and simulation fidelity (gravity vs. no-gravity). The ANOVA analysis revealed a significant main effect of simulation fidelity, F(1, 41) = 5.63, p = 0.023, η_p^2 = 0.132. Participants in the no-gravity condition (M = 5.69, SD = 0.75) rated the mean sensory fidelity scores significantly higher than participants in the gravity condition (M = 5.0, SD = 1.06). The ANOVA analysis did not reveal any significant main or interaction effects in the mean Involvement, Adaptation/Immersion and Interface Quality scores.

5.2.5 System Usability Results

We conducted a 2 × 2 multi-factorial ANOVA on the PSSUQ scores of participants by dimensions of System Usefulness, Information Quality, and Interface Quality. The independent between-subjects factors were presentation method (HMD vs. LSID) and simulation fidelity (gravity vs. no-gravity). With regards to System Usefulness, the ANOVA analysis revealed a significant main effect of presentation method, F(1, 41) = 3.350, p = 0.004, η_p^2 = 0.205. Participants in the HMD condition (M = 4.24, SD = 0.48) rated the system usefulness significantly higher than participants in the LSID condition (M = 3.67, SD = 0.65), p = 0.004. With regards to Information Quality and Interface Quality, the ANOVA analysis did not reveal any significant differences.

5.3 Performance Variables

Task performance variables in virtual reality interactions during spatial interaction for psychomotor skills training in precision metrology were extracted. The psychomotor variables extracted were time to complete (TTC) of each task within a module, accuracy of measurements via user answer difference, accuracy of measurement via physical difference, ability to interpret the Vernier scale via the reading difference, number of attempts in completing the task, and number of collisions during task performance. Each of the variables were treated with a 2 × 2 ANOVA, where the between-subjects factors were simulation fidelity (gravity or no-gravity) and presentation method (LSID or HMD). Post-hoc analyses were completed using Tukey HSD method.

5.3.1 Time to Complete

Mean time to complete overall across all the modules in the simulation based training were extracted and were treated with the 2 × 2 ANOVA analysis. The ANOVA analysis revealed a main effect of presentation method on mean TTC scores, F(1, 41) = 11.262, p = 0.002, η_p^2 = 0.23, and a presentation method by simulation fidelity interaction F(1, 41) = 4.136, p = 0.049, η_p^2 = 0.10. Pairwise comparisons revealed that participants in the LSID condition took significantly longer to complete the task when gravity was disabled (M = 103.6s, SD = 38.74) as compared to when gravity was enabled (M = 73.32s, SD = 12.34), p = 0.029 (see Figure 7).

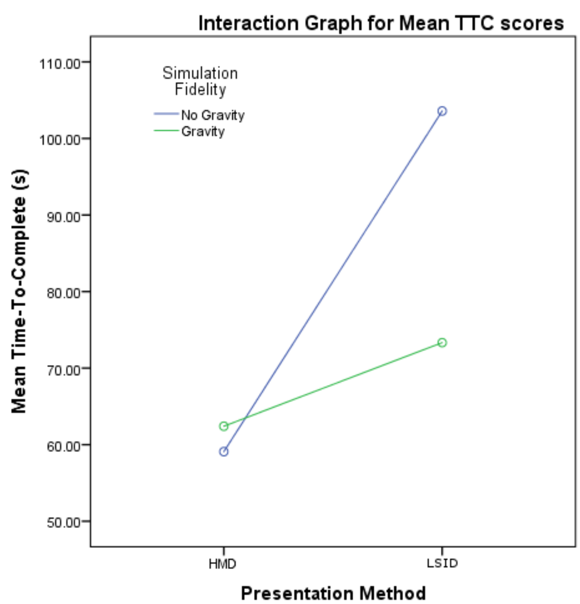

Figure 7: Interaction graph showing mean time-to-complete in presentation method by simulation fidelity interaction.

5.3.2 Reading Difference

The mean reading difference was a measure of the user's ability to read and interpret the Vernier scale, irrespective of the ground truth. The 2 × 2 ANOVA analysis revealed a significant main effect of presentation method on the overall mean reading difference scores, F(1, 41) = 4.71, $p = 0.036$, $\eta_p^2 = 0.11$. The overall mean reading difference was significantly higher in LSID condition (M = 0.033cm, SD = 0.018) as compared to the HMD condition (M = 0.023cm, SD = 0.012), $p = 0.036$. However, no significant effects were found for the physical difference or user answer difference performance metrics.

5.3.3 Number of Attempts

We compared the overall mean number of attempts taken by participants during the precision metrology tasks in the VR simulation via a 2 × 2 ANOVA. The ANOVA revealed a significant main effect of presentation method on the overall mean number of attempts, F(1, 41) = 5.76, $p = 0.022$, $\eta_p^2 = 0.135$. Overall, the mean number of attempts of participants in the LSID condition (M = 1.66, SD = 0.42) was significantly higher than participants in the HMD condition (M = 1.40, SD = 0.22), $p = 0.022$.

5.3.4 Number of Collisions

We compared the overall mean number of collisions during spatial interaction by participants while performing the precision metrology tasks in the VR simulation via a 2 × 2 ANOVA. The ANOVA analysis revealed a significant main effect of presentation method on the overall mean number of collisions, F(1, 41) = 9.98, $p = 0.003$, $\eta_p^2 = 0.213$. Overall, the mean number of collisions during spatial interactions in the VR simulation were significantly higher in the LSID condition (M = 334, SD = 121) as compared to the HMD condition (M = 196.55, SD = 149.35), $p = 0.022$.

5.3.5 Dominant and Non-Dominant Hand Distance

In order to study the distance covered by the dominant and non-dominant hand per task in the simulation, and if it differed by levels of presentation method and simulation fidelity, we conducted a 2 × 2 × 2 mixed model ANOVA analysis. The between subject variables were presentation method (HMD vs. LSID) and simulation fidelity (gravity vs. no-gravity), and the paired or related measures variable was the total distance covered by the dominant and non-dominant hand per task during the simulation. Pairwise post-hoc Tukey HSD analysis revealed that in the absence of gravity (M = 4.71m, SD = 4.01) the dominant hand covered a significantly larger distance than in the presence of gravity (M = 3.71m, SD = 3.05), $p < 0.001$ (see Figure 8).

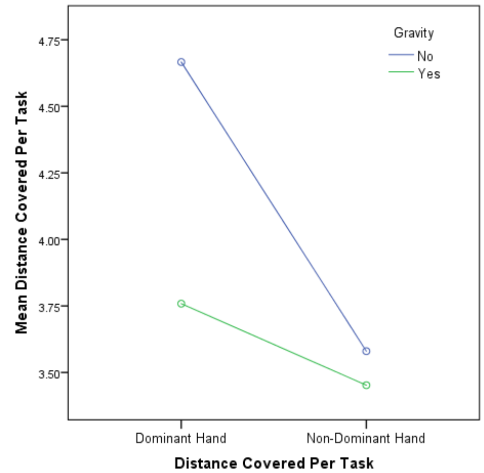

Figure 8: Distance covered by the dominant and non-dominant hand by the presence or absence of gravity

5.4 Qualitative Analysis

The After-Action Review tool was used to analyze participant behavior as they progressed through the training simulation. At the beginning of the Ruler module, participants generally interacted with only one hand and took measurements without picking up or positioning the other object at an advantageous orientation. But they quickly switched to using both hands either simultaneously or switching between them as they interacted with different objects in the scene. Although instructions were provided on how to interact with the instrument using the controllers, some participants tried to interact with movable parts of the instrument like the slide or the lock-screw using the other hand as one would in real life. This was only observed in the first module. While taking measurements, few participants used the 3D cursor to point at divisions on the scale to calculate the reading, analogous to using fingers in real life.

When gravity was present, most participants used both hands simultaneously to record their measurements. In the HMD condition, when participants found it difficult to align the object and the instrument together, they placed the object on the table in a stable fashion and took measurements using just one hand. This behavior was not observed in the LSID condition. An explanation for this could be that a top down view was possible in the HMD condition, making it easy to perform the above mentioned sequence of actions which the LSID condition did not afford.

In the absence of gravity, few participants used both hands simultaneously. Most used one hand at a time to move objects around and record measurements, thus employing a unimanual strategy. The majority of the participants switched to using just their dominant hand early on in the simulation to place objects at an advantageous position and measure them. Some switched hands depending on the object's position in the scene. Comparing the behavior between the different display conditions, participants positioned the instrument close to the camera (representative of their head position in the scene) and looked down for a better look in the HMD condition. Whereas, in the LSID condition, the participant moved the instrument as close as possible to them and leaned in towards the screen for a zoomed in view of the instrument.

Participants were asked for comments and suggestions via word of mouth during breaks and at the end of the experiment. Most participants reported that the task was fun and engaging and they preferred learning about instruments via the simulation. Some mentioned that the repetition of instructions on how to read the scale and use the instruments in each subsequent phase helped them while filling out post-surveys and in the skill-based performance assessment phase. Some participants who experienced the LSID condition mentioned it strained their eyes and was difficult to record measurements.

6 Discussion

Cognition questions in all levels of Bloom's Taxonomy revealed significant improvements from the pre to the post tests, however no significant differences were detected between the presentation method and simulation fidelity independent variables. These results were expected, nevertheless demonstrating the pedagogical benefits of the simulation is the first step in determining the usefulness of a virtual training application. Encouragingly, improvements in the higher levels of Bloom's Taxonomy demonstrated that the VR simulation can be effective for higher-order cognitive skills.

Participants rated the Performance and Mental Demand factors significantly higher in the HMD condition compared to the LSID condition in the NASA-TLX Workload Assessment. The Performance factor indicates how well users thought they completed the task and the Mental Demand score indicated the amount of attention they had to dedicate to the simulation. The significant Mental Demand scores can probably be attributed to users being completely immersed in the simulation, forcing them to fully concentrate on the task. HMD users also felt they were significantly more successful at accomplishing their task as compared to the LSID users due to their higher performance rating, suggesting that the HMD was easier to use than the LSID. HMD users reported significantly lower Frustration scores compared to the LSID condition. This was reinforced by statements from the users who said the HMD was easy and fun to use while some participants in the LSID condition stated they had difficulty adapting to the interaction metaphor.

Results from the subjective user questionnaires generally supported our first hypothesis in that users had a difficult time manipulating objects in the LSID condition, which in turn resulted in frustration, while metrics such as satisfaction, ease of use, and effectiveness were favored in the HMD condition despite this being the first time many participants used a head-mounted display. Performance variables such as the reading difference, number of attempts and number of collisions indicate that users had more difficulty interpreting the Vernier scale in the LSID condition. Incorrectly reading the Vernier scale forced the user to repeat their measurement and also significantly increased the number of contacts between the virtual objects since they usually had to re-measure the object.

The Sensory Fidelity factor of the Presence questionnaire revealed a significant main effect of gravity with participants rating the no-gravity condition better than the gravity condition. Questions in the Sensory Fidelity category involved the ability to examine objects closely. One key advantage of disabling gravity was that participants could bring the instrument close to their viewpoint and release it, allowing them to concentrate on interpreting the Vernier scale instead of having to continuously hold the instrument in addition to reading the scale. Furthermore, as the participant brought the instrument closer to them, their hands started to cross the boundary of the electromagnetic range of the Hydra, resulting in jitter. This was not a problem in the no-gravity condition since the user could simply freeze the object in place and move their head in closer for a better look in both presentation methods.

The majority of participants in the no-gravity condition adopted a unimanual approach to taking a measurement, in effect losing the dynamic frame of reference set by the non-dominant hand as described by Guiard [14]. The distance covered by the dominant hand was significantly higher than the non-dominant hand in the no-gravity condition, indicating that their dominant hand was doing most of the work. This strategy was also observed in the After-Action Review tool. The time to complete metric was significantly higher for participants in the LSID conditions without gravity, indicating that manipulating objects with one hand sequentially instead of in parallel had a significant negative effect on user performance, in turn lending support to our second hypothesis. However, in the real-world psychomotor assessment, we found no significant difference between presentation method or simulation fidelity on user scores and thus fail to reject our third hypothesis. This could be due to a number of factors including the fact that the task was not designed to be overly difficult for novice users and several participants had prior knowledge of how to use calipers. Results from the study were generally in line with the findings by McMahan et al. whereby higher fidelity simulations were preferable to mid-fidelity simulations if the end goal was to train a real-world task [25].

The presentation method between the HMD and LSID conditions contains several differences that could account for user performance. First, the HMD allows the user to rotate their head while the LSID's fish tank metaphor only allowed for position changes. Additionally, the user could physically move in for a closer look in the HMD condition more than the LSID afforded. The field-of-view was approximately 47° less in the LSID compared to the HMD. Additionally, the field-of-regard was 360° in the HMD while the LSID field-of-regard was approximately 60°. However, field-of-regard may have a low impact on performance due to the measurement task always being in front of the user and not requiring any navigation. The HMD's update rate was 75Hz and the LSID was 120Hz (60Hz per eye). The resolution of the LSID was 1920×1080 and the resolution for the HMD was 960×1080 per eye (1920×1080 total). A significant difference between the two methods is the visual motor offset between what the user's hands are doing and what they are perceiving on the LSID. This offset does not exist in the HMD condition because the user's hands are co-located between the physical and visual end-effectors. Moving objects closer than the minimum depth plane in LSID setups results in ocular discomfort when attempting to fuse a stereo image and thus limited the distance that virtual objects could be placed relative to the user's viewpoint. Each of these factors could have contributed to differences in performance, but this will require future work to properly evaluate each factor independently.

7 Conclusion and Future Work

With virtual training simulations becoming more prevalent across many application domains, there exists a growing need to determine what aspects of simulation and interaction fidelity are most beneficial for improving usability and learning outcomes. Conducting controlled studies allows researchers to identify the most appropriate aspects for training a real-world task in a virtual setting. The goal of this research was to examine the effects of varying interaction fidelity with respect to presentation method by comparing an HMD and immersive large-screen display. Results from this study suggest that users were less frustrated, more satisfied, performed faster and more accurately in the HMD condition compared to the LSID condition. Sometimes VR designers will break the laws of physics to enhance system usability at the cost of reducing simulation fidelity. This study examined the effect of enabling or disabling gravity to potentially enhance object manipulation and found that performance was degraded in the simplified no-gravity condition due to users adopting a unimanual approach to task completion. The results of this study suggest that VR designers should opt for higher fidelity interaction metaphors when possible. Future work will focus on more advanced training scenarios for automotive and aviation technicians that require hands-on training and examine the constituent components of each display method.

Acknowledgements

This work was supported by the National Science Foundation under Grant No. DUE-1104181. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

[1] F. Al-Zahrani. Web-Based Learning and Training for Virtual Metrology Lab. *Journal of Telecommunications*, 1(2):42–54, 2010.

[2] F. Argelaguet and C. Andujar. A survey of 3D object selection techniques for virtual environments. *Computers & Graphics*, 37(3):121–136, 2013.

[3] R. Balakrishnan and G. Kurtenbach. Exploring bimanual camera control and object manipulation in 3D graphics interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems the CHI is the limit - CHI '99*, number May, pages 56–62, New York, New York, USA, 1999. ACM Press.

[4] J. Bertrand, D. Brickler, S. Babu, K. Chalil Madathil, M. Zelaya, T. Wang, J. Wagner, A. Gramopadhye, and J. Luo. The role of dimensional symmetry on bimanual psychomotor skills education in immersive virtual environments. In *2015 IEEE Virtual Reality (VR)*, pages 3–10. IEEE, 2015.

[5] A. Bhatti, S. Nahavandi, Y. B. Khoo, D. Creighton, J. Anticev, and M. Zhou. Haptically enable interactive virtual assembly training system development and evaluation. In *SIMTECT 2009. Proceedings of the 2009 International Design Engineering Technical Conferences & Computers and Computers and Information in Engineering Conference*, pages 1–9, 2009.
[6] B. Bloom. *Taxonomy of educational objectives - the classification of educational goals - Handbook I: cognitive domain*. David McKay Company, Inc, 1956.
[7] D. Bowman, E. Kruijiff, J. J. LaViola Jr., and I. Poupyrev. *3D User Interfaces: Theory and Practice*. Addison Wesley, 2004.
[8] W. Buxton and B. Myers. A study in two-handed input. *ACM SIGCHI Bulletin*, 17(4):321–326, 1986.
[9] S. R. Chaiken, P. C. Kyllonen, and W. C. Tirre. Organization and Components of Psychomotor Ability. *Cognitive Psychology*, 40(3):198–226, 2000.
[10] R. H. Dave. *Developing and writing behavioural objectives*. Educational Innovators Press, 1970.
[11] S. Gebhardt, S. Pick, F. Leithold, B. Hentschel, and T. Kuhlen. Extended Pie Menus for Immersive Virtual Environments. *IEEE Transactions on Visualization and Computer Graphics*, 19(4):644–651, 2013.
[12] T. Y. Grechkin, T. D. Nguyen, J. M. Plumert, J. F. Cremer, and J. K. Kearney. How does presentation method and measurement protocol affect distance estimation in real and virtual environments? *ACM Transactions on Applied Perception*, 7(4):1–18, 2010.
[13] M. W. Gribnau and J. M. Hennessey. Comparing single- and two-handed 3D input for a 3D object assembly task. In *CHI 98 conference summary on Human factors in computing systems - CHI '98*, number April, pages 233–234, New York, New York, USA, 1998. ACM Press.
[14] Y. Guiard. Asymmetric Division of Labor in Human Skilled Bimanual Action. *Journal of Motor Behavior*, 19(4):486–517, 1987.
[15] S. G. Hart and L. E. Staveland. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in psychology*, pages 139–183. 1988.
[16] K. Hinckley, R. Pausch, J. C. Goble, and N. F. Kassell. A survey of design issues in spatial input. In *Proceedings of the 7th annual ACM symposium on User interface software and technology - UIST '94*, pages 213–222, New York, New York, USA, 1994. ACM Press.
[17] K. Hinckley, R. Pausch, J. C. Goble, and N. F. Kassell. Passive real-world interface props for neurosurgical visualization. In *Proceedings of the SIGCHI conference on Human factors in computing systems celebrating interdependence - CHI '94*, volume 30, pages 452–458, New York, New York, USA, 1994. ACM Press.
[18] K. Hinckley, R. Pausch, and D. Proffitt. Attention and visual feedback: The bimanual frame of reference. In *Proceedings of the 1997 symposium on Interactive 3D graphics - SI3D '97*, pages 121–ff., New York, New York, USA, 1997. ACM Press.
[19] K. Hinckley, R. Pausch, D. Proffitt, and N. F. Kassell. Two-handed virtual manipulation. *ACM Transactions on Computer-Human Interaction*, 5(3):260–302, 1998.
[20] A. Leganchuk, S. Zhai, and W. Buxton. Manual and cognitive benefits of two-handed input: an experimental study. *ACM Transactions on Computer-Human Interaction*, 5(4):326–359, 1998.
[21] K. S. Lehmann, J. P. Ritz, H. Maass, H. K. Cakmak, U. G. Kuehnapfel, C. T. Germer, G. Bretthauer, and H. J. Buhr. A Prospective Randomized Study to Test the Transfer of Basic Psychomotor Skills From Virtual Reality to Physical Reality in a Comparable Training Setting. *Annals of Surgery*, 241(3):442–449, 2005.
[22] J. R. Lewis. IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, 7(1):57–78, 1995.
[23] R. W. Lindeman, J. L. Sibert, and J. K. Hahn. Towards usable VR. In *Proceedings of the SIGCHI conference on Human factors in computing systems the CHI is the limit - CHI '99*, pages 64–71, New York, New York, USA, 1999. ACM Press.
[24] K. Chalil Madathil, K. Frady, A. K. Gramopadhye, J. Bertrand, and R. S. Hartley. Integrating case studies in an online asynchronous learning environment: An empirical study to evaluate the effect on community college student learning outcomes and engagement when case studies are presented using virtual reality. *ASEE Annual Conference and Exposition, Conference Proceedings*, 2016-June, 2016.
[25] R. McMahan. *Exploring the Effects of Higher-Fidelity Display and Interaction for Virtual Reality Games*. Doctoral dissertation, Virginia Polytechnic Institute and State University, 2011.
[26] M. R. Mine, F. P. Brooks, and C. H. Sequin. Moving objects in space. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques - SIGGRAPH '97*, pages 19–26, New York, New York, USA, 1997. ACM Press.
[27] J. Murayama, L. Bougrila, and Y. Luo. SPIDAR G&G: A two-handed haptic interface for bimanual VR interaction. In *Proceedings of Eurohaptics*, pages 138–146, 2004.
[28] M. Nabiyouni, A. Saktheeswaran, D. A. Bowman, and A. Karanth. Comparing the Performance of Natural , Semi-Natural , and Non-Natural. pages 243–244, 2015.
[29] M. H. Oermann. Psychomotor Skill Development. *The Journal of Continuing Education in Nursing*, 21(5):202, 1990.
[30] D. Parmar, J. Bertrand, S. V. Babu, K. Chalil Madathil, M. Zelaya, T. Wang, J. Wagner, A. K. Gramopadhye, and K. Frady. A comparative evaluation of viewing metaphors on psychophysical skills education in an interactive virtual environment. *Virtual Reality*, 20(3):141–157, 2016.
[31] D. Parmar, J. Bertrand, B. Shannon, S. V. Babu, K. Chalil Madathil, M. Zelaya, Tianwei Wang, J. Wagner, K. Frady, and A. K. Gramopadhye. Interactive breadboard activity simulation (IBAS) for psychomotor skills education in electrical circuitry. In *2014 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 181–182. IEEE, 2014.
[32] E. D. Ragan, R. Kopper, P. Schuchardt, and D. A. Bowman. Studying the Effects of Stereo, Head Tracking, and Field of Regard on a Small-Scale Spatial Judgment Task. *IEEE Transactions on Visualization and Computer Graphics*, 19(5):886–896, 2013.
[33] F. D. Rose, E. A. Attree, B. M. Brooks, D. M. Parslow, and P. R. Penn. Training in virtual environments: transfer to real world tasks and equivalence to real task training. *Ergonomics*, 43(4):494–511, 2000.
[34] U. Schultheis, J. Jerald, F. Toledo, A. Yoganandan, and P. Mlyniec. Comparison of a two-handed interface to a wand interface and a mouse interface for fundamental 3D tasks. In *2012 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 117–124. IEEE, 2012.
[35] M. Slater, M. Usoh, and A. Steed. Depth of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 3(2):130–144, 1994.
[36] B. G. Witmer, C. J. Jerome, and M. J. Singer. The Factor Structure of the Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, 14(3):298–312, 2005.