Comparative Evaluation of Digital Writing and Art in Real and Immersive Virtual Environments

Chi-Hsuan Hsu*
National Chiao Tung University

Chi-Han Chung†
National Chiao Tung University

Rohith Venkatakrishnan‡
Clemson University

Roshan Venkatakrishnan§
Clemson University

Yu-Shuen Wang¶
National Chiao Tung University

Sabarish V. Babu‖
Clemson University

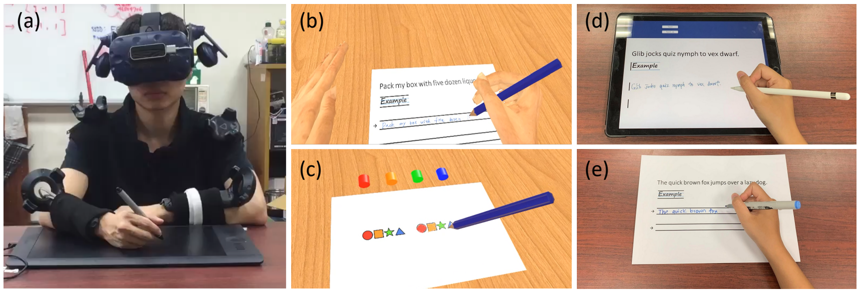

Figure 1: (a) Participant in our simulation, (b) digital writing in VR with self-avatar, (c) coloring in VR without self-avatar, (d) digital writing in the real world, and (e) writing in the real world using a pen and paper.

2021 IEEE Virtual Reality and 3D User Interfaces (VR) | 978-1-6654-1838-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/VR50410.2021.00089

ABSTRACT

Virtual reality (VR) experiences currently tend to focus on full body interactions. However, fine motor control in actions such as writing and drawing are seldom studied. Challenges include the inability to perceive fine details due to the low resolution of head mounted displays, the difficulty in simulating fine motor actions in virtual environments, tracking instabilities, latency issues, etc. State of the art VR has managed to address a host of such concerns, supporting a variety of input mechanisms for activities such as writing, sketching, immersive modeling, etc. With VR increasingly being applied in education and medical contexts where writing and note taking is a crucial, it is important to study how well humans can perform these tasks in VR. In a between-subjects empirical evaluation, we studied participants' fine motor coordination with several digital input based writing and artistic tasks performed both in virtual and real world settings, further examining the effects of providing a virtual self avatar on task performance. We integrated multiple tracking systems and applied inverse kinematics to animate the virtual body and simulate hand motions. We went on to compare how different the outputs of these digital input metaphors are to a real world pen and paper approach in an effort to ascertain where we currently stand in being able to support writing and note taking in virtual world contexts. Overall, it seems to be the case that while writing and artistic activities can be successfully supported in VR applications using specialized input devices, the accuracy with which users perform such tasks is significantly higher in the real world, highlighting the need for developments that support such fine motor tasks in VR.

Keywords: Writing and Art in Virtual Reality, Perception-Action Coordination, Fine Motor Control

1 INTRODUCTION

Virtual reality (VR) as a technology has exploded in recent years with the increasing availability of affordable display hardware, tracking and interaction devices. This growth in the popularity of VR has led to its application to a variety of areas like gaming [17], training [24], therapy [45], education [7], etc. As a consequence, there seems to be an increasing prevalence of educational VR applications that foster learning in medical [42], STEM [35] and other fields. In such contexts, note taking has been shown to be a highly significant and integral part of learning [19]. This makes it important for virtual experiences to afford users the ability to jot down notes, sketch diagrams, etc. as a part of the interaction paradigm associated with the medium [12, 37].

While it is possible to facilitate note taking via keyboard based text entry and illustrations via mouse-based interaction, these interaction metaphors cannot be easily integrated into immersive virtual environments (IVEs) due to challenges in tracking and rendering associated with these interaction devices, and our ability to convincingly simulate fine motor actions on them [29]. It has also been shown that users seek and prefer interaction metaphors that resemble real world scenarios over traditional mouse-based interaction [6, 18]. To address such concerns, natural interaction metaphors have been proposed for note taking and drawing using tracked stylus and tablet surfaces to allow users to perform the fine motor task of recording observations in an intuitive manner. This metaphor supports users drawing upon perceptuo-motor coordination acquired from real world writing, further providing tactile augmentation between the tip of the stylus and virtual canvas [7, 15, 37]. Such natural metaphors have also been shown to promote a greater sense of presence, enhance educational benefits and improve the efficacy of task performance in VR [7, 9]. While latency issues between the motor component and the corresponding content to be rendered on the virtual canvas have proven to be problematic and detrimental to user performance in the past [37], contemporary VR systems have managed to overcome these issues with advancements in the technology. Through high resolution displays, hybrid tracking techniques, high

---

*e-mail: ace810408@gmail.com
†e-mail: chihhichihhi@gmail.com
‡e-mail: rohithv@g.clemson.edu
§e-mail: rvenkat@g.clemson.edu
¶e-mail: yushuen@cs.nctu.edu.tw
‖e-mail: sbabu@clemson.edu

2642-5254/21/$31.00 ©2021 IEEE
DOI 10.1109/VR50410.2021.00089

frame-rate tracking and rendering, state of the art VR now supports natural pen based interactions for writing and art-work rather easily, Google's Tilt Brush being one such popular example [25].

Handwriting involves complex visual-perceptual-motor processing leveraging perception, motor, and executive functions. These functions are synchronized and integrated at various levels to produce a word [11, 39]. It is an inherently bimanual task wherein a division of labor between the dominant and non-dominant hand is created, requiring the brain to simultaneously control movements from both hands [22]. Several studies have examined how to lay out interface elements to afford unimanual and bimanual division of labor of fine motor actions [5], some involving writing and drawing on a digital canvas [8]. As such, while writing is accomplished primarily by the dominant hand, the non dominant hand may aid in detecting, analyzing, and integrating visual perceptual information, such as paper angles or spatial boundaries, relative to dominant hand movements [22, 46]. Research has underscored the importance of virtual self-representation via self-avatars in perception-action coordination tasks [27, 31, 34], demonstrating the importance of visual, kinesthetic and proprioceptive feedback [5, 13]. The view of the self-body serves as an egocentric ruler to perceive and scale the users' action capabilities and affords better motor control [4, 10]. Self avatars have also been shown to aid users' cognitive processes and reduce mental workload in letter recall tasks performed in VR [41]. Body scaled self-embodiment via self-avatars is becoming commonplace in VR simulations, due to an increase in the amount of affordable tracking hardware, robust processors, and readily accessible inverse kinematics based animation frameworks. Given that writing, tracing, drawing all require fine dexterous motor control within the near-field or personal space [20, 43], it is postulated that the presence of a body-scaled articulated self-avatar may afford scaling of action capabilities and facilitate better motor control in a similar manner to real world interactions.

In this work, we empirically evaluated how the fine motor perceptual tasks of writing, tracing and coloring are performed in real and virtual world settings involving digital inputs on a fixed size digital canvas. We further examined how the affordance of a body scaled self-avatar affects the performance of writing and art activities in the virtual world. We also discuss how these digital input interaction metaphors differ from a traditional pen and paper paradigm, analyzing the outputs using objective image processing techniques as well as subjective similarity assessments.

2 RELATED WORK

There have been several studies that have looked into perceptuomotor coupled behaviors involved in writing and drawing in the real world. Handwriting is considered to be a highly specialized motor task. Thomassen and van Galen have pointed out that fluent, cursive handwriting is a fine distal task using the most delicate muscles of the hand and fingers, the required precision for which is accurately developed much later than locomotion, reaching and grasping [43]. It is a specialized motor task that requires coordination of form production and accurate spatial judgment, utilizing stored motor knowledge and accurate ordering, spacing and lineation. Gowen et al. examined eye-hand coordination in tracing and drawing of abstract objects in the real world [20]. They found that during tracing, the pen tip and eye were tightly coupled, with participants making a series of small saccades just in front of the moving pen. During drawing, they found that saccades were fewer and larger and that pursuit was less frequent. Guiard presented a kinematic chain as a model for human intermanual division of labor in skilled activities such as drawing and drafting [22]. The model postulates that the two end effectors represent two manual motors that cooperate with one another, where the non-dominant hand serves as a reference in fine motor tasks for the dominant hand. Motions produced by the dominant hand tend to articulate with motions produced by the non-dominant hand.

The literature with respect to digital writing features contributions ranging from virtual typing, effects of virtual reality viewing in perception-action coordination, to interaction metaphors for typing, sketching and drawing in 3D space. The earliest work with respect to writing in virtual worlds, conducted by Poupyrev et al., presented a virtual notepad that supported two handed interaction with tactile augmentation, resulting in a highly intuitive tablet-stylus interaction writing paradigm in VR [37]. It was found that while this interaction metaphor leveraged experience of real world writing, meeting users' expectations was challenging due to technical issues with the latency in visualizing writing output on the digital canvas. Research has also examined the benefits of using a pen and touch mechanism for bimanual input on horizontal interfaces, where it was found that users were faster and committed fewer errors using the pen and touch input as compared to either touch only or pen only inputs [6]. More recently, Pham and Stuerzlinger compared the use of a 3D pen (stylus) to the virtual controller for selection of targets in VR and found that the pen outperformed the controller in fine motor target selection [36]. This result suggests that a virtual pen might be better suited for the fine motor task of writing or drawing than a controller.

A number of works have looked into the importance of providing self representations for tasks requiring fine motor control actions. Grubert et al. examined the effects of typing between no hand, animated hand, fingertip visualization, and a video view of typing in real world on typing performance in VR [21]. Although they found no difference on text entry speed, but found that errors were significantly lower in the fingertip visualization and video integration of real world typing, as compared to the no hand and animated hand condition [21]. They noted that the fidelity of tracking of participants' hands in the fine motor action of typing may be responsible for the poor performance in VR, as compared to the fingertip visualization and video inlay conditions. When performing fine motor pick and place tasks, Argelaguet et al. found that the sense of agency was less in the virtual hand representation that rendered in an abstract manner, but the sense of body ownership was greater in the virtual hand representation that was most realistic [1]. Furthermore the authors of [38] highlighted the importance of gender matching avatars in promoting a sense of agency in fine motor tasks performed in virtual worlds. When it comes to bimanual fine motor tasks such as mechanical skills and surgical training, virtual reality based training has been shown to improve psychomotor skills learning as compared to traditional learning methods in the presence of self avatars [5, 26].

A number of works have investigated drawing and sketching in extended reality (XR) environments. The authors of [44], conceptualized a bimanual input technique for AR modeling that combines a standard smartphone with a 3D-printed pen used for modeling objects mid-air. It was found that users prefer casting a ray through the pen tip for the selection and translation of objects. This work involved tasks based on the pen being used for mid air selection and translation and didn't comprehensively probe into evaluating how well the conceptualized interaction metaphor fared on drawing or sketching performance. With respect to the drawing process, dynamic elements such as the order of compilation, speed, length, and pressure of strokes are considered important as they may reveal the technique and process of the artist in question. Research by Fernenado et al. suggests that sensing, visualizing and sharing these aspects of the creative process could help shape art making and viewing experiences [16]. Other researchers have explored the immersive 3D modeling paradigm analyzing the behaviours of users with different spatial abilities while drawing in VR, finding that spatial abilities affect the shape of the drawings but not the precision of lines [3]. While the majority of the work that has looked into immersive modeling has focused on mid air based modeling and sketching, there have been some investigations that adopt a more real world like sketching scenario. Work on this front has compared traditional sketching on a physical surface to sketching in VR, with and without a physical surface to rest the stylus on, finding that the lack of a physical drawing surface is a major cause of inaccuracies in VR drawing [2].

In summary, we found that there is little or no research that compares and contrasts perceptual-motor task performance in writing and art activities such as handwriting, sketching, coloring and tracing in VR and RW experiences, let alone examining how the affordance of self avatar representations affect the performance of such activities in contemporary VR experiences.

3 STUDY DESIGN

Towards empirically evaluating how fine motor perceptual tasks like writing, tracing and coloring are performed in real and virtual world settings involving digital inputs, and further examining how the affordance of a body scaled self-avatar affects the performance of these activities in the virtual world, we conducted a between subjects study with three cell block conditions, each involving participants performing the fine motor perceptual tasks described in section 5.3. The conditions featured the manipulations of interest and we refer to them as *VREI*: VR with an embodied self-avatar, *VRI*: VR without self-avatar, and *RWI*: real world digital writing. The *VREI* and the *VRI* conditions involved VR as a medium to perform the tasks whereas the *RWI* condition involved participants performing the tasks on tablet in the real world. These three conditions hence featured digital inputs. In the *VREI* condition, participants were afforded a gender-matched, body-scaled self avatar in the virtual environment and hence had to perform the tasks in the presence of a self avatar. In contrast, the *VRI* condition did not feature a self avatar and involved participants performing the tasks without a self avatar. The study consisted of two sessions, the first of which involved participants experiencing one of the three digital input conditions. The second session involved participants performing the same tasks in a real world pen and paper scenario which we call *PAPER*. The second session was conducted after the completion of the first session and it was ensured that participants did not feel fatigued or cybersick before they could proceed to engage in the second session. The size and location of the paper matched those of the virtual canvas used in the VR conditions to ensure consistency between conditions. The pen and paper paradigm serves as a benchmark against which the digital interfaces can be evaluated.

3.1 Research Questions and Hypotheses

In this work, we investigated the following research questions:

1. To what extent can users perceive and coordinate their actions in the process of writing, tracing and coloring in IVEs as compared digital input in the real world?

2. How does the affordance of self-avatars affect users' performance in writing, tracing and coloring in IVEs?

3. To what extent is participants' performance in digital writing, tracing and coloring tasks, similar to a traditional pen and paper based paradigm?

Our hypotheses were as follows:

- H1 - Participants' performance will be less erroneous in real world digital input than in immersive virtual environments.

- H2 - Participants' performance in immersive virtual environments will be less erroneous in the presence of a co-located self-avatar as compared to using a stylus only input.

- H3 - Participants' performance in the paper and pen session will be similar to that of the real world digital input, but different than the performance in the immersive virtual environment.

Figure 2: Body and finger proportion measurements used for calibration of self-avatar.

4 SYSTEM DESCRIPTION

4.1 VR and RW Systems

To support handwritten digital inputs for the tasks, we built an interface using the Unity3D engine. This interface was used both in the real world and VR to ensure consistency between conditions featuring digital inputs. On top of this, we built a virtual scene for the VR conditions of this study (VRI and VREI) in which participants performed the tasks using the aforementioned interface. This interface was deployed on an IPad Pro (10.5-inch screen) for the RWI condition and leveraged the Apple pencil as a stylus. The interface presented participants with the tasks one after another, recording strokes and logging data on the time spent by the participants to accomplish each task. For the PAPER paradigm, we printed each task on a paper and these were presented to the participants. Upon completion, these papers were digitally scanned to compare against the other conditions. The time spent for the tasks in PAPER session was manually recorded. Given that the experiment featured both digital and traditional mediums (i.e. PAPER), a meaningful analysis required that we measure the errors in the physical coordinate system (pixel area error) to analyze these results in a vigorous, objective fashion. Apropos this, the writing space in all the conditions was set to 25.3×15.8 cm; and the stylus nib was set to 0.05 cm wide.

To provide haptic feedback for participants in the VR conditions, we set up a desk and a chair in the real world, placing a Wacom tablet on the desk. Participants were seated on the chair and could perform the tasks using the Wacom stylus. The physical desk, chair, and tablet were tracked using HTC Vive tackers, ensuring the positional and orientational co-location in the real world and VR. In addition to this, we tracked participants' upper body using HTC Vive puck trackers (see figure 3), using this to render a gender-matched, body-scaled self avatar for the VREI condition. This also allowed us to track their movements throughout the experiment. A dynamic virtual rendering of the Wacom stylus was achieved using a combination of two systems. The magnetic sensors on the tablet allowed for the detection of position and orientation of the stylus; and the system that tracked the hands and wrist was used in conjunction to virtually render the stylus. In the VRI condition, there was no self avatar and the virtual stylus alone was rendered. Participants were equipped with the trackers in all conditions to ensure consistency between conditions. The only difference between the two VR conditions was the presence and absence of a self-avatar. This setting allowed us to specifically evaluate the effects of embodied interaction in VR, further allowing for the comparison of these fine motor actions against real world settings.

4.2 Body Tracking and Gesture Animation

Prior to the experiment, we measured the body proportions of each participant and adjusted the self-avatar properties including the size, scale and location of the hands and fingers based on these mea-

Figure 3: (Left) Participants equipped with HTC Vive controllers and trackers.(Top right) Controller strapped to arm and used in conjunction with inverse kinematics model to determine elbow joint position.(Bottom right) Plastic finger cushion on stylus for resting thumb.

sured proportions. Figure 2 depicts the joints that were used as measurement references. We attached the HTC Vive controllers to the participants' forearms using straps, attaching trackers on their shoulders (Figure 3), using these along with the coordinates of the head mounted display to accurately reconstruct the pose of the upper body. We set the positions and the orientations of the key joints as constraints and applied inverse kinematics to determine the remaining joint positions.

Tracking gestures is a challenging task. Although there have been many computer vision techniques proposed to solve this problem, they all assume that users' hands are not in contact. When participants are holding a stylus, tracked gestures can be easily wrong and unstable. We hence manually created a personalized stylus-holding grip for each subject's dominant hand based on how they held the stylus. To achieve this, participants held the stylus with their dominant hand based on which a self avatar was calibrated. We measured the relative positions from the stylus nib to the second knuckle of participants' thumbs and index fingers, and then applied the positions to fine tune the self-avatar's finger positions. These relative positions helped in animating participants' dominant hands for which we first detected the pen nib's position using the Wacom tablet's tracking system, then back tracked the knuckles' positions, followed by applying the corresponding computed positions to animate the dominant hand using inverse kinematics. This strategy not only achieved a natural writing animation but also reduced the stylus-hand intersections without additional collision detection. In addition, we stuck a plastic finger cushion on the stylus, as illustrated in Figure 3 to further help the participants position their thumb and index finger on the stylus. Thus, without having to remove the HMD, participants could stick to their grip via passive haptic feedback received as a result of the cushion on the stylus tip. For the non-dominant hand, we modeled a flat palm gesture, which matched participants' non-dominant hands. The end result was that the modeled and animated self-avatar with the dominant and non-dominant hands matched the participants' hand poses, making for a highly realistic 'writing in VR' experience.

4.3 Stylus Tracking

We tracked the stylus using two different systems. When the stylus is close to the tablet (i.e., within 1.5 cm), its position and orientation can simply be detected from the Wacom system's electromagnetic sensors. The Wacom tablet has a robust 6DoF tracking of the stylus nib when close to the tablet, and can track the nib's pose and force upon contact with the surface. This allows for high precision tracking that befits situations involving fine motor interactions associated with writing and art in VR. When the stylus was relatively further away from the tablet, we used participants' tracked wrists to compute where the stylus was and render it accordingly. We seamlessly transitioned between these two systems, making for a realistic rendering of the stylus throughout the course of the simulation.

5 EXPERIMENT

5.1 Participants

We recruited a total 45 participants for this Institution Review Board (IRB) approved study using advertisements on social media. Their ages ranged from 21 to 40 (M=24.18, SD=3.34). All participants had normal or corrected to 20/20 vision. To reduce interpersonal differences, we recruited only right-handed users. Participants with less than 6 hours of sleep prior to the experiment were excluded from taking part in the study. Any individuals that indicated feelings of fatigue and body pain were also excluded. After exclusion, we still had a total of 45 participants, 23 of which were female, the remaining, male (22).

5.2 Procedure

Upon arrival at the laboratory, participants were greeted and asked to read and sign a consent form (informed consent). After having consented to participate in the study, participants filled out a demographics questionnaire that included information about their backgrounds and experience with VR, 3D movies and video games. On completing these surveys, participants were briefed about the experiment and were randomly assigned to one of the three conditions with digital input (VREI, VRI and RWI). Following this, participants' body segment proportions were measured as illustrated in Figure 2, and used to calibrate a personalized body-scaled, gender matched self-avatar for participants in the VREI condition. Participants then held the stylus based on which their personalized grip of the pen was calibrated (See Section 4.2). This was performed even for participants in the VRI condition to ensure the accurate rendering of the stylus in the virtual world. Participants in the VR conditions underwent a body-ownership acclimation phase in which they stood against a mirror in the virtual world. Even though the VRI condition did not feature a self avatar, the mirror allowed for participants to acclimate to the tracked stylus' movements. Participants that were given self avatars performed egocentric and exocentric pointing tasks wherein they touched their arms, bodies and faces several times to acclimate to the self avatar, and pointed to different objects in the virtual scene. The virtual environment featured several objects such as a big mirror, a clock, a plant, etc., allowing for an immersive experience. To ensure visibility in the scene, participants were asked to report the time displayed on the clock. Following this, participants performed the writing and art (tracing and coloring) tasks the details of which are described in Section 5.3. The order of the tasks were randomized. Participants were allowed to take a break whenever they felt fatigued. Upon completion of the tasks, participants filled out the IBM System Usability Questionnaire [30], NASA TLX Questionnaire [23] and the SUS presence questionnaire [40]. A similar procedure was adopted for participants in the RWI condition except that they were not immersed in VR. All participants had to finally perform these tasks using a traditional pen and paper matching the canvas size, position, and stylus geometry of the conditions involving digital inputs, thus constituting the PAPER condition of the study. Lastly, all participants engaged in semi structured interviews that gave us feedback about their preferences and perceptions of the experience as a whole. On average, the study took an hour to complete (no significant differences between conditions).

5.3 Tasks

We designed multiple tasks-2 writing, 1 tracing, and 1 coloring task. These tasks were chosen because they leverage high degrees of perception-action coordination, while being representative of real world educational settings. The tasks required participants to perceive the location of the canvas, their virtual end effectors, the location of the stylus nib in creating digital writing and art. Participants had to perform all tasks regardless of the condition they

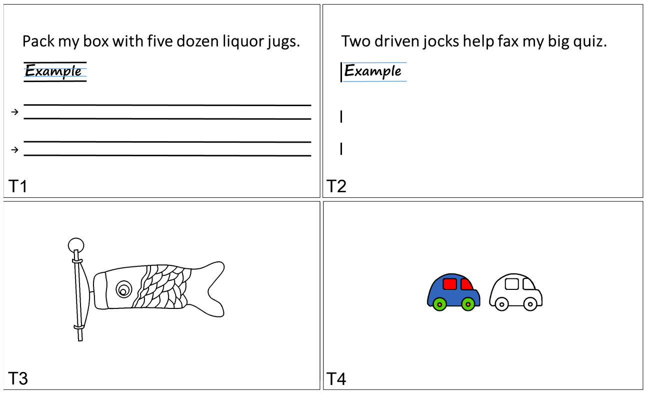

Figure 4: (T1 and T2) writing tasks; (T3 and T4)tracing and coloring tasks respectively.

were assigned to. To compare their perceptuo-motor capabilities against a real world scenario with pen and paper, we intentionally avoided provisioning the ‘undo’ and ‘erase’ functions. For every task, participants were provided with a practice trial before the testing trials commenced. The practice trial performance was not analyzed. We describe the details of each task in this section. See Figure 4.

T1 - Writing within boundary: Participants were shown a sentence on the top of the writing area, and were asked to rewrite the sentence in a bounding area specified by two horizontal lines. There were a total of 5 sentences that had to be written under 3 different height constraints. This made a total of 15 sentences, the order of which was randomized. Participants had to fulfil the height constraint of writing within the two horizontal lines and could use additional lines if they ran out of space

T2 - Writing without boundary: This task in general was very similar to the first. However, the height constraint was indicated by a vertical bar on the left rather than two horizontal lines. Therefore, the participants had to imagine the area by themselves when writing, thus requiring higher perceptuo-motor coordination.

T3 - Tracing: Users were shown a black contour image in the middle of the drawing area and were asked to trace the image. There were 5 images in the task (Fish, Kawaii, Turtle, Umbrella and House), the order of which was randomized. They were instructed to trace all the contours as precisely as they could, avoiding white areas.

T4 - Coloring: Participants were shown a colored reference image on the left and a contour image on the right, respectively. They were asked to color the image on the right and make it resemble the reference image to the highest degree possible. A total of 5 images was used in this task (mushroom, car, elephant, dragonfly and flower), the order of which was randomized. Notice that colors required in different regions are quite distinct. This design choice reduced the probability of selecting wrong colors.

5.4 Results

The quantitative data gathered from the objective and subjective variables were analyzed using a mixed model repeated measures, Analysis of Variance (ANOVA) analysis. The between-subjects independent factor was the conditions featuring digital inputs (RWI, VRI, VREI). The within-subjects repeated measures variable was the sub-tasks within a category of tasks such as writing within a boundary of 1.5cm, 1.25cm and 1.0 cm, as well as the digital input and the real world paper and pen based writing and drawing that participants completed in two within-subjects sessions of our experiment. Where possible, we have objectively and subjectively analyzed the participants’ performance in the virtual and real world digital writing and drawing in a between-subjects manner, and compared the digital input with the real world paper and pen based writing and drawing in a repeated measures manner.

On all quantitative data, parametric ANOVA analyses were conducted on the data after carefully verifying that the underlying assumptions were met – namely the data in the samples were normally distributed and error variance between samples were equivalent. We ensured that Box’s test of equality of covariance matrix was not significant. Levene’s test was conducted to verify homogeneity of variance, and Mauchly’s test of sphericity was conducted to ensure that the error variance in groups of samples was equivalent. Pairwise post-hoc tests between levels of the between-subjects variables was conducted using Tukey’s HSD analysis, and between levels of the within-subjects variables was conducted using the Bonferroni adjusted alpha method.

5.4.1 Writing Task - Within Boundary

Pixel Area: The number of pixels of error that exceeded the double line boundary of writing was converted to total pixel error area (mm^2) and subsequently analyzed between the conditions of real and virtual world digital input, and compared against pen and paper based writing, across trials involving 1.5, 1.25 and 1.0 cm line widths. The mean error area across trials of writing within the boundary were analyzed using a 3 (conditions – RWI vs. VRI vs. VREI) × 3 (line width – 1.5 vs. 1.25 vs. 1.0cm) × 2 (digital input vs. pen and paper) mixed model ANOVA analysis. The ANOVA analysis revealed a significant main effect of line width on mean pixel error area, F(2, 84) = 8.74, $p < 0.001$, $\eta^2 = 0.17$. The mean pixel error area in line width 1.50cm trials (M = 0.023, SD = 0.058) was significantly lower than in 1.25cm (M = 0.068, SD = 0.16) $p = 0.042$, and 1.00cm (M = 0.112, SD = 0.21) trials $p < 0.001$. No other main or interaction effects were found.

5.4.2 Writing Task - Without Boundary

Pixel Area: The number of pixels in digital writing that exceeded the width of the reference bar was converted to an error pixel area and was subsequently analyzed between the conditions of real world and VR digital input, and compared against pen and paper based writing, across trials involving 1.5, 1.25 and 1.0 cm bar widths. The mean error pixel area (mm^2) across trials of writing without the boundary line were analyzed using a 3 (conditions) × 3 (bar width - 1.5 vs. 1.25 vs. 1.0cm) × 2 (digital input vs. pen and paper) mixed model ANOVA analysis. The ANOVA analysis revealed a significant main effect of reference bar width on mean error pixel area, $F(2, 84) = 45.77$, $p < 0.001$, $\eta^2 = 0.52$. Overall, the mean error pixel area were significantly lower in the bar width 1.50cm trials (M = 2.41, SD = 3.6), as compared to bar width 1.25cm (M = 4.27, SD = 5.58) $p < 0.001$, and 1.0cm (M = 7.10, SD = 8.12) trials $p < 0.001$. Additionally, mean error pixel area were significantly lower in bar width 1.25cm trials as compared to bar width 1.0cm trials $p < 0.001$. No other significant main or interaction effects were found.

Subjective Rating of Participants’ Writing Quality: We recruited four human raters who were blinded to the conditions to rate on a scale of 1 (least similar) through 10 (most similar), the participants’ writing quality in each of the digital input session (RWI, VRI and VREI) to that of the same participants’ paper and pen output in the real-world baseline writing session. The human raters provided similarity scores of writing in both the writing within boundary as well as the writing without boundary trials, in line or bar widths of 1.50, 1.25 and 1.00 cms in all conditions. In each type of trial (writing within boundary or writing without boundary), the subjective rating scores were analyzed by a 3 (trial type) × 3 (condition-RWI, VRI or VREI) repeated measures mixed model ANOVA. The ANOVA analysis of the participants’ scores in writing within boundary did not reveal any significant effects.

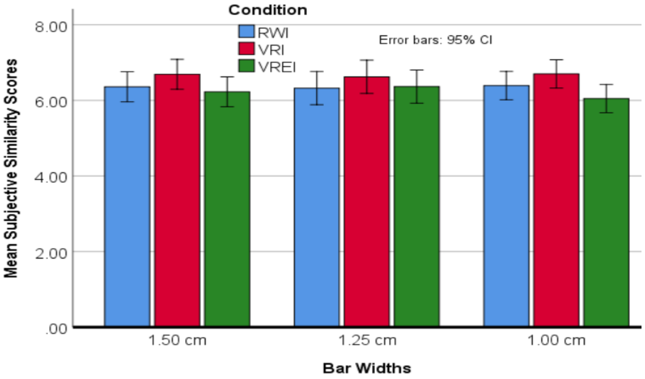

Figure 5: Mean subjective similarity scores of participants' writing without boundary between digital input and paper and pen session.

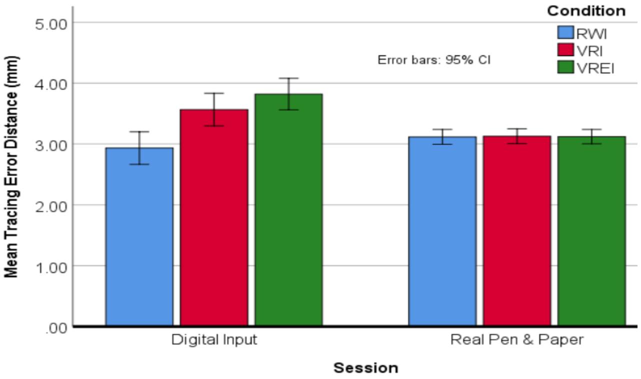

Figure 6: Mean error distance in session by condition interaction.

Table 1: Post-hoc comparisons on subjective similarity scores for the tracing task in the digital input conditions. note: * denotes p <.05, ** denotes p <.01 ,*** denotes p <.001

	Fish		Kawaii	
	Means	SD	Means	SD
VREI	6.33	1.35	6.05	1.11
VRI	6.33	1.35	6.3	1.46
RWI	7.5	0.85	7.73	0.70
Bonferroni	RWI >	VRI** VREI**	RWI >	VRI** VREI**

However, the ANOVA analysis of participants' writing without boundary revealed a significant condition by bar width interaction effect on subjective rating scores, $F(4, 84) = 2.83$, $p = 0.030$, $\eta^2 = 0.12$ (Figure 5). The Intraclass Correlations (ICC) of interrater reliability was ICC(2,4)=0.72. Post-hoc pairwise analysis revealed that in the writing without boundary in 1.00cm bar width trials, participants' similarity scores in VRI condition (M = 6.7, SD = 0.59) was significantly higher than participants in the VREI condition (M = 6.04, SD = 0.88) with that of their real world pen and paper writing, $p = 0.044$.

5.4.3 Tracing Task

In a manner similar to the writing tasks, the canvas resolution and drawing line pixel width were standardized across the conditions and sessions of the experiment, so that a reliable and meaningful comparison of error in performance can be generated objectively. Carefully designed and implemented image processing algorithms extracted the number of error pixels traced on the contour of the figures outlines, and from that the perpendicular error pixel distance from the figure contour was computed and as an objective performance variable. These were subjected to a 3 (Conditions) × 5 (Objects) × 2 (Sessions) mixed model repeated measures Analysis of Variance (ANOVA) analysis, after the underlying assumptions of the parametric test were verified.

Error Distance: The ANOVA analysis on the mean error distance in the tracing task revealed a significant main effect of object $F(4, 160) = 29.33$, $p < 0.001$, $\eta^2 = 0.42$, a significantly main effect of condition $F(2, 40) = 6.93$, $p = 0.003$, $\eta^2 = 0.26$, a significant main effect of session $F(1, 160) = 27.68$, $p < 0.001$, $\eta^2 = 0.41$, a significant session by condition interaction $F(2, 160) = 18.79$, $p < 0.001$, $\eta^2 = 0.48$, a significant objects by session interaction $F(4, 160) = 4.067$, $p = 0.004$, $\eta^2 = 0.092$, and a significant objects by session by condition 3-way interaction effect $F(8, 160) = 2.05$, $p = 0.04$, $\eta^2 = 0.093$. Among the interaction effects, we focus on the session by condition interaction in post-hoc analysis as the interaction involving condition is the most interesting to our study, also having the relatively highest effect size (Figure 6).

Post-hoc pairwise analysis using Bonferroni method revealed no significant difference in the mean error distance in the RWI condition between the digital input and real world pen and paper condition. However, in the VRI condition mean error distance was significantly higher in the digital input session (M = 3.56, SD = 0.70) as compared to the pen and paper session (M = 3.13, SD = 0.28), $p = 0.008$. Also in the VREI condition mean error distance was significantly higher in the digital input session (M=3.8, SD=0.56) as compared to the pen and paper session (M = 3.12, SD = 0.28), $p < 0.001$. Post-hoc Tukey's HSD pairwise comparisons did not reveal any significant differences between RWI, VRI and VREI on mean error distance in the pen and paper session tracing performance. However, Tukey's HSD pairwise comparisons revealed that the mean tracing error distance in the digital input session was significantly higher in VRI (M = 3.56, SD = 0.65) as compared to RWI (M = 2.93, SD = 0.29) condition $p = 0.005$, and significantly higher in VREI (M = 3.82, SD = 0.47) as compared to RWI condition $p < 0.001$.

Subjective Similarity Rating of Tracing: Similar to the subjective similarity score generation in the writing task, four human raters provided similarity scores of tracing quality between the real world session and the digital input session for each of the objects that varied in complexity in all conditions. The Intraclass Correlations (ICC) of interrater reliability was ICC(2,4)=0.68. The subjective rating scores were analyzed by a 5 objects × 3 (condition) repeated measures mixed model ANOVA. The analysis revealed a significant main effect of objects F(4, 168) = 4.40, $p = 0.002$, $\eta^2 = 0.095$, a significant main effect of condition $F(2, 42) = 17.87$, $p < 0.001$, $\eta^2 = 0.46$, and a significant objects by condition interaction effect $F(8, 168) = 2.43$, $p = 0.016$, $\eta^2 = 0.11$ (Figure 7). Overall, subjective similarity scores for RWI were the highest, followed by VRI and were the lowest in the VREI conditions. Post-hoc pairwise comparisons using Tukey's HSD analyses were conducted on subjective similarity scores on the tracing task and some of the significant differences are summarized in table 1

5.4.4 Coloring Task

Digital image processing techniques were used to detect the number of color pixels incorrectly placed outside the region, or missing

Table 2: Post-hoc comparisons on pixel error area for coloring task in the digital input conditions. note: * denotes p <.05, ** denotes p <.01 ,*** denotes p <.001

	Mushroom		Elephant		Flower	
	Mean	SD	Mean	SD	Mean	SD
VREI	12.68	10.58	30.15	27.0	17.22	10.96
VRI	12.38	11.98	27.05	24.72	13.09	11.96
RWI	2.95	3.75	5.62	6.56	2.61	2.73
Tukey's HSD	RWI <	VRI* VREI*	RWI <	VRI* VREI**	RWI <	VRI** VREI***

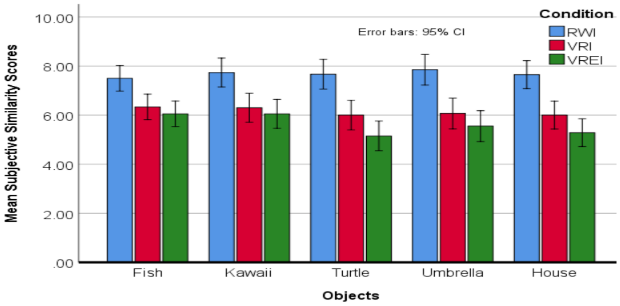

Figure 7: Mean subjective similarity scores of tracing task performance between digital input conditions by objects interaction.

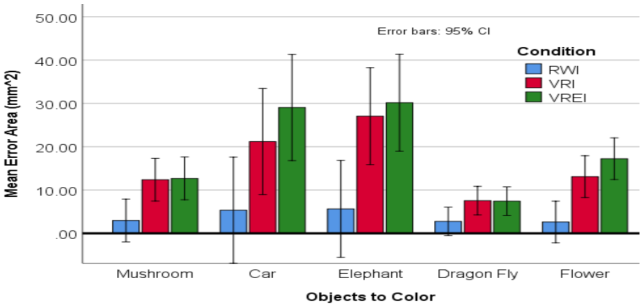

Figure 8: Mean error area of coloring task for condition by objects interaction.

pixels of color within a region. The total number of error pixels per image was converted to the total pixel error area and proportion of total pixels in an image that are colored in error, as objective metrics. Due to technical challenges, we were unable to extract objective error metrics of the participants' coloring performance in the scanned digital images of the pen and paper session. In the objective metrics measured, we statistically compared the error scores in a 5 (object colored) × 3 (condition) ANOVA analysis, ensuring that all the assumptions of the parametric ANOVA analysis were met.

Error Area: The ANOVA analysis on the mean error area (mm^2) of coloring pixels revealed a significant main effect of objects $F(4, 168) = 16.68$, $p < 0.001$, $\eta^2 = 0.28$, a significant main effect of condition $F(2, 42) = 6.31$, $p = 0.004$, $\eta^2 = 0.23$, and a significant objects by condition interaction effect $F(8, 168) = 2.905$, $p = 0.005$, $\eta^2 = 0.12$ (Figure 8). Post-hoc pairwise comparisons using Tukey's HSD analyses were conducted on pixel error area and significant differences are summarized in table 2.

Subjective Evaluation of Coloring Task: We statistically compared the mean similarity scores of the four ratings, between the digital input conditions (RWI, VRI and VREI). The Intraclass Correlations (ICC) of interrater reliability was ICC(2,4)=0.71. The mean coloring similarity scores for five different coloring tasks (Mushroom, Car, Elephant, Dragon Fly, and Flower) that varied in complexity of the size, scale and geometric complexity of the regions to be colored distinctly were compared using a 3 (condition) × 5 (objects) mixed model ANOVA analysis. The within-subjects factor was the objects drawn, and the between-subjects factor was the conditions. The ANOVA analysis on the subjective evaluation of coloring task revealed a significant main effect of objects $F(4, 168) = 4.48$, $p = 0.002$, $\eta^2 = 0.096$, a significant main effect of condition $F(2, 42) = 9.66$, $p < 0.001$, $\eta^2 = 0.31$, and a significant objects by condition interaction effect $F(8, 168) = 3.45$, $p = 0.002$, $\eta^2 = 0.14$ (Figure 9). Post-hoc pairwise comparisons using Tukey's HSD analyses were conducted on the subjective similarity scores and significant differences are summarized in table 3.

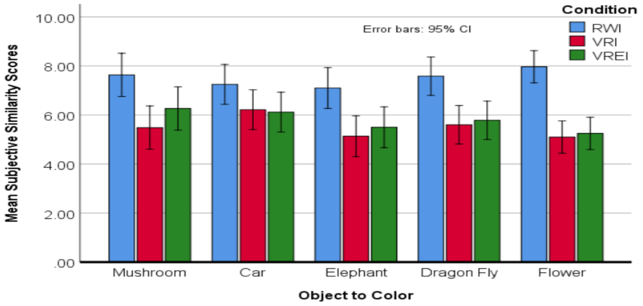

Figure 9: Subjective similarity scores of digital coloring in various conditions with the participants' paper and pen coloring performance.

Table 3: Post-hoc comparisons on subjective similarity scores for the coloring task in the digital input conditions. note: * denotes p <.05, ** denotes p <.01 ,*** denotes p <.001

	Dragon Fly		Elephant		Flower	
	Mean	SD	Mean	SD	Mean	SD
VREI	5.78	1.19	5.5	1.54	5.25	0.84
VRI	5.6	1.67	5.13	1.69	5.10	1.53
RWI	7.58	1.60	7.1	1.54	7.98	1.31
Tukey's HSD	RWI >	VRI** VREI**	RWI >	VRI** VREI*	RWI >	VRI*** VREI***

5.4.5 Workload and other Measures

Amongst all dimensions of workload obtained from the NASA TLX, the ANOVA analysis for frustration was significant $F(2, 44) = 5.59$, $p = 0.007$. Post-hoc pairwise Tukey HSD comparisons revealed that mean frustration scores were significantly higher in the VREI condition (M=5.27, SD=2.40) as compared to VRI (M=2.87, SD=2.2) $p = 0.012$ and RWI (M=3.07, SD=1.91) $p = 0.023$ conditions. Additionally, there were no significant differences found in the aspects of presence [40] and system usability [30] between the conditions.

5.4.6 Qualitative Feedback

Users in the VR conditions were highly appreciative of the accuracy of the rendered stylus, indicating that it felt naturally co-located with the physical one. Several participants commented that writing without a boundary required them to frequently look back at the reference bar, dynamically adjusting their writing positions. Participants in the VREI conditions mentioned that their self avatar representations occluded the canvas, requiring them to move their heads in order to complete the the tasks successfully. However, this issue wasn't raised by participants in the other conditions due to the absence of a virtual self avatar. Even though the self-avatar hand in the VREI condition was calibrated to match the proportions of subjects' real hands, animation limitations due to inverse kinematics and other idiosyncrasies may have contributed to the perception of occlusion. A few users in the VR conditions were unsure if pixels close to boundaries were colored or had the right color, requiring them to move their head close to the canvas to confirm the same.

6 DISCUSSION

When participants wrote on the digital canvas under the guidance of the twin lines, the results showed that they were able to complete the tasks with the different types of line spacing in a roughly equivalent manner. Understandably, the mean errors were highest when the spacing between the lines was small (1.0cm) and lowest when the spacing was larger (1.5cm). It was also observed that the real world digital writing condition and the pen and paper paradigm produced similar performances. The writing task that did not include boundary lines was designed to be more difficult, requiring one to perceive the height of the reference bar and accordingly coordinate their writing motion to satisfy the bar's height constraint. We were surprised to find the same pattern of results in this task as we did in the previous. Despite the relatively impoverished viewing and motor control of the VR world, users were able to write in a manner similar to how they would on pen and paper even when the reference bar was only 1cm tall. When we analyzed the subjective similarity rating of the participants' writing between their pen and paper and digital input sessions, we surprisingly found that the writing was statistically more similar in the VR condition in the absence of self-embodiment than in the presence of the same in the writing task T2, deviating from results obtained by Knierim et. al [28] who found that typists benefited from seeing their hands in VR. These differences in results may be attributed to the paradigm of text entry employed in the virtual world with [28] employing a physical keyboard as opposed to the hand based writing paradigm employed in this study. We expected that participants' writing performance would be better in the presence of hand and body scaled self-avatars, as they could scale their actions by using their body as a ruler in a manner similar to other situations in VR such as reaching [14], locomotion [33], and perceiving threats [1]. However, with writing, we actually found objective performance to be similar in the VR and real world digital input conditions, and similarity with paper and pen writing to be higher in VR without self-embodiment.

The tracing task required users to perceive and follow curved contours of very thin lines on a digital canvas. It was found that both VR conditions were associated with higher errors than the RWI condition for this task. Surprisingly, users in the VREI condition produced significantly higher errors than those in VRI condition. With respect to the subjective similarity ratings, a sort of similar trend was observed with the ratings in the RWI condition's output being most similar to the PAPER session's output followed by the VRI and VREI conditions respectively. From these results, it is clear that for tracing tasks, VR based digital input schemes achieved using an integrated tablet are not as effective as digital inputs supported in the real world. More work is needed to investigate why exactly this is the case. In the coloring task, participants had to perceive the location, size and scale of the different regions to be colored, filling those regions with the appropriate colors using the stylus. For all objects, the errors produced in the VR conditions were significantly higher than those produced in the RWI condition. This again goes to show that VR based digital inputs are not as good as digital mediums used in the real world. Similar to the tracing task, we obtained results suggesting that digital inputs in the real world achieved using tablets produce significantly better results than those used in VR systems, in terms of how similar they are to coloring performed on pen and paper. Interestingly, in some of the coloring trials, similarity scores of participants' coloring performance with their pen and paper session was higher in VR with self-embodiment than VR without.

While no differences were found between the conditions on aspects of presence and system usability, we did find that levels of frustration was significantly higher in the condition that afforded users with a self avatar in VR. This is likely due to the occlusion issues that users alluded to in the interviews. Overall, users preferred performing the tasks in the real world, probably due to the hardware and tracking limitations associated with VR. Compared to medium-field body based interactions such as travel and selection [32], in which the self-body avatar plays a positive role in the interaction, the tracking requirement for near-field fine motor control are much higher [5]. Since users performed tasks involving their virtual end effectors, it may be that they were able to perceive small scale inconsistencies between their physical and virtual bodies.

In summary, we found that the first hypothesis was supported in our study in that participants' performance in writing, tracing, and coloring was less erroneous in the real world digital input condition as compared to the VR conditions. We also found that participants' performance in the RWI condition was most similar to that of the pen and paper session with the VR conditions generating mixed results. In the writing and tracing tasks, the subjective similarity scores were higher when there was only a stylus as compared to the VR condition with a self-avatar, but was opposite in the coloring task, partially supporting hypothesis three. Contrary to our expectations, the VR condition with the self-avatar generated the highest number of errors in performance as compared to the VR with stylus only condition, and thus the second hypothesis was not supported. A limitation of our study was that the self-avatar hand gestures were generated using a carefully calibrated inverse kinematic model leveraging the HTC Vive tracker at the wrist, the hand dimension measurements, and the stylus's tracked position which may have been potentially insufficient. While the hand size, color and shape were carefully calibrated to match users' real hands, we do not know to what extent a higher fidelity self-avatar hand and finger tracked gesture system, leveraging superior motion capture technologies, could improve the perception-action coordination involved in writing, tracing and coloring tasks in VR. It is also worth noting that the styli used in the different conditions (Wacom and Apple Stylus) are not identical in properties like shape and weight, possibly influencing these results.

7 CONCLUSION AND FUTURE WORK

In this work, we empirically examined the performance and perceptual-motor coordination involved in writing and artistic tasks in the real world and VR. We further evaluated how the affordance of a body scaled self-avatar affects the performance of these activities (writing, tracing and coloring) in VR, leveraging inverse kinematics and other computational schemes to support animations of the self avatar and its end effectors. We went on to compare how different the outputs of these digital writing metaphors are to a real world pen and paper approach in an effort to ascertain where we currently stand in being able to support writing and art in VR. In a two session between subjects study, participants first performed tasks in one of three digital input conditions: 1) *VREI*: VR with an embodied self-avatar, *VRI*: VR without self-avatar, and *RWI*: real world digital input on a tablet. In the second session, participants performed the same tasks on pen and paper. The results of our study showed that in general, digital input metaphors used in the real world are superior to a stylus-tablet based digital input metaphor used in VR, producing lesser errors and better resembling outputs obtained on the traditional pen and paper paradigm. In terms of writing, participants in the VR conditions were more or less able to match the performance of those in the RWI condition, showing promise for developers aiming to use VR in contexts involving handwriting. Surprisingly, we observed that the affordance a self avatar did not improve the performance on these tasks. We supported embodiment and gesture tracking using an inverse kinematics approach with low cost commodity tracking hardware. Our results suggest that it may still have been insufficient to enhance perceptuo-motor coordination as compared to a stylus only interaction in VR. Further research is needed to investigate how the presence of a high fidelity self-avatar hands may potentially enhance handwriting and artwork in VR scenarios.

In future work, we aim to investigate how advanced hand tracking technologies and high fidelity animation frameworks for self avatars affect the perceptual motor coordination required for these tasks. Specifically, we are interested in employing haptic gloves and a motion capture based animation framework towards investigating if these tracking and animation techniques affect users' performances on the same tasks conducted in this study. Our long term interests lie in developing novel techniques to improve the efficacy of writing, art and other fine motor perceptual tasks in VR.

ACKNOWLEDGMENTS

This work was partly supported by a grant from the Taiwan Ministry of Science and Technology (07-2811-E-009-004). We would like to thank all the participants for devoting their time to take part in this study.

REFERENCES

[1] F. Argelaguet, L. Hoyet, M. Trico, and A. Lécuyer. The role of interaction in virtual embodiment: Effects of the virtual hand representation. In *Virtual Reality (VR), 2016 IEEE*, pp. 3–10. IEEE, 2016.

[2] R. Arora, R. H. Kazi, F. Anderson, T. Grossman, K. Singh, and G. W. Fitzmaurice. Experimental evaluation of sketching on surfaces in vr. In *CHI*, vol. 17, pp. 5643–5654, 2017.

[3] M. D. Barrera Machuca, W. Stuerzlinger, and P. Asente. The effect of spatial ability on immersive 3d drawing. In *Proceedings of the 2019 on Creativity and Cognition*, pp. 173–186. 2019.

[4] A. Berti and F. Frassinetti. When far becomes near: Remapping of space by tool use. *Journal of cognitive neuroscience*, 12(3):415–420, 2000.

[5] A. Bhargava, J. W. Bertrand, A. K. Gramopadhye, K. C. Madathil, and S. V. Babu. Evaluating multiple levels of an interaction fidelity continuum on performance and learning in near-field training simulations. *IEEE transactions on visualization and computer graphics*, 24(4):1418–1427, 2018.

[6] P. Brandl, C. Forlines, D. Wigdor, M. Haller, and C. Shen. Combining and measuring the benefits of bimanual pen and direct-touch interaction on horizontal interfaces. In *Proceedings of the working conference on Advanced visual interfaces*, pp. 154–161. ACM, 2008.

[7] Y.-T. Chen, C.-H. Hsu, C.-H. Chung, Y.-S. Wang, and S. V. Babu. ivrnote: Design, creation and evaluation of an interactive note-taking interface for study and reflection in vr learning environments. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 172–180. IEEE, 2019.

[8] L. D. Cutler, B. Fröhlich, and P. Hanrahan. Two-handed direct manipulation on the responsive workbench. In *Proceedings of the 1997 symposium on Interactive 3D graphics*, pp. 107–114. ACM, 1997.

[9] K. A. Darabkh, F. H. Alturk, and S. Z. Sweidan. Vrcdea-tcs: 3d virtual reality cooperative drawing educational application with textual chatting system. *Computer Applications in Engineering Education*, 26(5):1677–1698, 2018.

[10] B. Day, E. Ebrahimi, L. S. Hartman, C. C. Pagano, and S. V. Babu. Calibration to tool use during visually-guided reaching. *Acta psychologica*, 181:27–39, 2017.

[11] B. de Haas and G. Rees. Multiple stages of cross-modal integration in visual processing. comment on" crossmodal influences on visual perception" by l. shams & r. kim. *Physics of life reviews*, 7(3):287, 2010.

[12] M. F. Deering. Holosketch: a virtual reality sketching/animation tool. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 2(3):220–238, 1995.

[13] E. Ebrahimi, B. Altenhoff, L. Hartman, J. A. Jones, S. V. Babu, C. C. Pagano, and T. A. Davis. Effects of visual and proprioceptive information in visuo-motor calibration during a closed-loop physical reach task in immersive virtual environments. In *Proceedings of the ACM Symposium on Applied Perception*, pp. 103–110. ACM, 2014.

[14] E. Ebrahimi, L. S. Hartman, A. Robb, C. C. Pagano, and S. V. Babu. Investigating the effects of anthropomorphic fidelity of self-avatars on near field depth perception in immersive virtual environments. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1–8. IEEE, 2018.

[15] R. Elmgren. Handwriting in vr as a text input method, 2017.

[16] P. Fernando, J. Weiler, and S. Kuznetsov. A rough sketch of the freehand drawing process. 2019.

[17] D. L. Fisher, A. Pollatsek, and A. Pradhan. Can novice drivers be trained to scan for information that will reduce their likelihood of a crash? *Injury Prevention*, 12(suppl 1):i25–i29, 2006.

[18] C. Forlines, D. Wigdor, C. Shen, and R. Balakrishnan. Direct-touch vs. mouse input for tabletop displays. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 647–656. ACM, 2007.

[19] L. Ganske. Note-taking: A significant and integral part of learning environments. *ECTJ*, 29(3):155–175, 1981.

[20] E. Gowen and R. C. Miall. Eye–hand interactions in tracing and drawing tasks. *Human movement science*, 25(4-5):568–585, 2006.

[21] J. Grubert, L. Witzani, E. Ofek, M. Pahud, M. Kranz, and P. O. Kristensson. Effects of hand representations for typing in virtual reality. *arXiv preprint arXiv:1802.00613*, 2018.

[22] Y. Guiard. Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *Journal of motor behavior*, 19(4):486–517, 1987.

[23] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, vol. 52, pp. 139–183. Elsevier, 1988.

[24] W. L. Heinrichs, P. Youngblood, P. M. Harter, and P. Dev. Simulation for team training and assessment: case studies of online training with virtual worlds. *World journal of surgery*, 32(2):161–170, 2008.

[25] G. INC. Tilt Brush by Google web description, 2018.

[26] J. Jordan, A. Gallagher, J. McGuigan, and N. McClure. Virtual reality training leads to faster adaptation to the novel psychomotor restrictions encountered by laparoscopic surgeons. *Surgical Endoscopy*, 15(10):1080–1084, 2001.

[27] E. Jun, J. K. Stefanucci, S. H. Creem-Regehr, M. N. Geuss, and W. B. Thompson. Big foot: Using the size of a virtual foot to scale gap width. *ACM Transactions on Applied Perception (TAP)*, 12(4):16, 2015.

[28] P. Knierim, V. Schwind, A. M. Feit, F. Nieuwenhuizen, and N. Henze. Physical keyboards in virtual reality: Analysis of typing performance and effects of avatar hands. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–9, 2018.

[29] J. J. LaViola Jr, E. Kruijff, R. P. McMahan, D. Bowman, and I. P. Poupyrev. *3D user interfaces: theory and practice*. Addison-Wesley Professional, 2017.

[30] 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.

[31] Q. Lin, J. Rieser, and B. Bodenheimer. Stepping over and ducking under: The influence of an avatar on locomotion in an hmd-based immersive virtual environment. In *Proceedings of the ACM Symposium on Applied Perception*, pp. 7–10. ACM, 2012.

[32] R. P. McMahan, D. A. Bowman, D. J. Zielinski, and R. B. Brady. Evaluating display fidelity and interaction fidelity in a virtual reality game. *IEEE Transactions on Visualization & Computer Graphics*, (4):626–633, 2012.

[33] E. A. McManus, B. Bodenheimer, S. Streuber, S. De La Rosa, H. H. Bülthoff, and B. J. Mohler. The influence of avatar (self and character) animations on distance estimation, object interaction and locomotion in immersive virtual environments. In *Proceedings of the ACM SIGGRAPH Symposium on applied perception in graphics and visualization*, pp. 37–44. ACM, 2011.

[34] B. J. Mohler, S. H. Creem-Regehr, W. B. Thompson, and H. H. Bülthoff. The effect of viewing a self-avatar on distance judgments in an hmd-based virtual environment. *Presence: Teleoperators and Virtual Environments*, 19(3):230–242, 2010.

[35] D. Parmar, J. Isaac, S. V. Babu, N. D'Souza, A. E. Leonard, S. Jörg, K. Gundersen, and S. B. Daily. Programming moves: Design and evaluation of applying embodied interaction in virtual environments to enhance computational thinking in middle school students. In *Virtual Reality (VR), 2016 IEEE*, pp. 131–140. IEEE, 2016.

[36] D.-M. Pham and W. Stuerzlinger. Is the pen mightier than the controller? a comparison of input devices for selection in virtual and augmented reality. In *25th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–11, 2019.

[37] I. Poupyrev, N. Tomokazu, and S. Weghorst. Virtual notepad: handwriting in immersive vr. In *Proceedings. IEEE 1998 Virtual Reality Annual International Symposium (Cat. No. 98CB36180)*, pp. 126–132. IEEE, 1998.

[38] V. Schwind, P. Knierim, C. Tasci, P. Franczak, N. Haas, and N. Henze. " these are not my hands!" effect of gender on the perception of avatar hands in virtual reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 1577–1582, 2017.

[39] L. Shams and R. Kim. Bayesian priors and multisensory integration at multiple levels of visual processing: Reply to comments on "crossmodal influences on visual perception". *Physics of Life reviews*, 7(3):295–298, 2010.

[40] M. Slater, M. Usoh, and A. Steed. Depth of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 3(2):130–144,

1994.

[41] A. Steed, Y. Pan, F. Zisch, and W. Steptoe. The impact of a self-avatar on cognitive load in immersive virtual reality. In *2016 IEEE Virtual Reality (VR)*, pp. 67–76. IEEE, 2016.

[42] A. Stevens, J. Hernandez, K. Johnsen, R. Dickerson, A. Raij, C. Harrison, M. DiPietro, B. Allen, R. Ferdig, S. Foti, et al. The use of virtual patients to teach medical students history taking and communication skills. *The American Journal of Surgery*, 191(6):806–811, 2006.

[43] A. J. Thomassen and G. P. van Galen. Handwriting as a motor task: Experimentation, modelling, and simulation. In *Advances in psychology*, vol. 84, pp. 113–144. Elsevier, 1992.

[44] P. Wacker, O. Nowak, S. Voelker, and J. Borchers. Arpen: Mid-air object manipulation techniques for a bimanual ar system with pen & smartphone. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2019.

[45] J. Wald and S. Taylor. Efficacy of virtual reality exposure therapy to treat driving phobia: a case report. *Journal of behavior therapy and experimental psychiatry*, 31(3-4):249–257, 2000.

[46] F. R. Wilson. *The hand: How its use shapes the brain, language, and human culture*. Vintage, 1999.