Evaluation of Point-to-Point Reaching Performance in Mixed Reality and Virtual Reality

Tyler G. Petrie[1,2,4], Chen Zhao[1] and Michael J. Fu[1,2,3]

Case Western Reserve University,

2123 Martin Luther King Jr. Dr.,

Glennan Bldg. Rm. 321, Cleveland OH, 44106, USA[1]

MetroHealth System, Cleveland, OH, USA[2]

Department of Veterans Affairs Northeast Ohio Healthcare System,

Cleveland, OH, USA[3]

tyler.petrie@case.edu[4]

March 12, 2024

Abstract

Recent immersive mixed reality (MR) and virtual reality (VR) displays enable users to use their hands to interact with both veridical and virtual environments simultaneously. Therefore, it becomes important to understand the performance of human hand-reaching movement in MR. Studies have shown that different virtual environment visualization modalities can affect point-to-point reaching performance using a stylus, but it is not yet known if these effects translate to direct human-hand interactions in mixed reality. This paper focuses on evaluating human point-to-point motor performance in MR and VR for both finger-pointing and cup-placement tasks. Six performance measures relevant to haptic interface design were measured for both tasks under several different visualization conditions ("MR with indicator", "MR without indicator", and "VR") to determine what factors contribute to hand-reaching performance. A key finding was evidence of a trade-off between reaching motion confidence" measures (indicated by throughput, number of corrective movements, and peak velocity) and "accuracy" measures (indicated by end-point error and initial movement error). Specifically, we observed that participants tended to be more confident in the "MR without Indicator" condition for finger-pointing tasks. These results contribute critical knowledge to inform the design of VR/MR interfaces based on the application's user performance requirements.

1 Introduction

Recent commercially available virtual reality (VR) and mixed reality (MR) devices use immersive head-mounted displays and hand gesture recognition to enable human operators to use their hands to interact with virtual objects. Although many studies have evaluated the effect of immersive VR display modalities on hand-reaching performance (M. Fu, Hershberger, Sano, & Çavuşoğlu, 2012; Khademi, Mousavi Hondori, Dodakian, Cramer, & Lopes, 2013; Ha & Woo, 2010), much less is known about how MR display modalities impact hand-reaching performance.

Specifically, it is not well understood how different visualizations of the hand in VR and MR affect human reaching performance. Since immersive head-mounted VR displays visually obstruct the veridical environment, operators must rely on computer-generated indicators of the hand as they interact with the virtual environment. In contrast, immersive head-mounted MR displays have transparent lenses that allow operators to see their veridical hand co-located with virtual objects, so it is unclear if computer-generated hand indicators are necessary. MR development guidelines recommend using fingertip indicators to provide visual cues related to when the operator's hand has intersected user interface elements like menus (Microsoft Corp., 2019). Additionally, finger indicators have been demonstrated as a potential solution for depth perception mismatch that can occur when the veridical hand does not visually occlude virtual objects in MR (Tang, Hu, Fu, & Cohen-Or, 2020). However, the effect of fingertip indicators on reaching movement performance has not been systematically examined.

Our study focuses on evaluating human point-to-point reaching performance during finger-pointing and cup-placement tasks under various MR and VR experiment conditions, including the presence of fingertip indicators. Point-to-point reaching is an important component not only of human-computer interaction, but also of computer-assisted rehabilitation for people with movement impairment related to neurological injuries such as stroke, spinal cord injury, and cerebral palsy. Finger-pointing is relevant to rehabilitation because it requires shoulder-elbow reaching, which is necessary to place the hand about a person's functional workspace in preparation for activities of daily living. Point-to-point reaching while grasping a cup is also important because it requires grasp, relocation, and object placement movements that are fundamental to a person's ability to feed themselves. Both finger-pointing and cup-placement are commonly trained in rehabilitation clinics and in VR-based training.

Virtual environments are promising for rehabilitation because they allow repetitive motor skill practice (which is important for relearning limb control) in an engaging, reproducible, and safe way (Scalona, Hayes, Del Prete, Palermo, & Rossi, 2019; M. J. Fu, Knutson, & Chae, 2015). VR-based rehabilitation can also motivate rehabilitation participants by providing performance feedback and by simulating salient training tasks using objects of daily living, such as handles, utensils, and cups (Holden, 2005; M. J. Fu et al., 2015). MR-based rehabilitation may provide people with disabilities the additional benefit of practicing hand movements using veridical objects of daily living that have mass and volume, which is absent from VR. MR's ability to project computer-generated images onto a person's veridical environment (such as their own home) may also ease the transfer of rehabilitation exercises into everyday life (Pridmore, Cobb, Hilton, Green, & Eastgate, 2007) and better engage the user physically and mentally (Duff et al., 2010).

Community and rehabilitation use of MR can benefit from head-mounted MR displays, which highlights the importance of understanding motor performance in this context. We know from studies on non-head-mounted VR displays that point-to-point reaching performance can be impacted by viewpoint orientation and co-location (M. Fu et al., 2012). Others have also reported that Fitts' Task performance can be affected by the type of veridical prop used to represent virtual objects in the context of a non-colocated augmented reality display using a computer monitor (Ha & Woo, 2010). However, studies comparing point-to-point reaching performance between co-located immersive, head-mounted MR and VR display modalities are lacking in the literature.

1.1 Study Objectives and Hypotheses

This work investigated how VR and MR visualization conditions affect human performance on a 3D version of Fitts' task in the contexts of finger-pointing and cup-placement. There are 3 visualization conditions: 1) MR with indicators where a virtual fingertip pointer or 3D cup model was superimposed upon their veridical counterparts, 2) MR without indicators where no fingertip or cup overlays are displayed, requiring interaction between the veridical finger or cup and virtual targets, and 3) VR with indicators where we shrouded the HoloLens to block the veridical environment from the participant's view, showing only virtual object interactions.

Objective 1: Compare finger-pointing task performance between MR and VR environments. *Hypothesis 1:* We hypothesize that performance will be increased during the finger-pointing tasks under both MR visualization experiment conditions with than in the without indicator condition and VR environment. *Objective 2:* Comparing cup-placement task performance between MR and VR environments. *Hypothesis 2:* We hypothesize that performance will be increased during the cup-placement tasks under both MR visualization experiment conditions (with and without indicator) than in the VR environment.

2 Methods

2.1 Participants

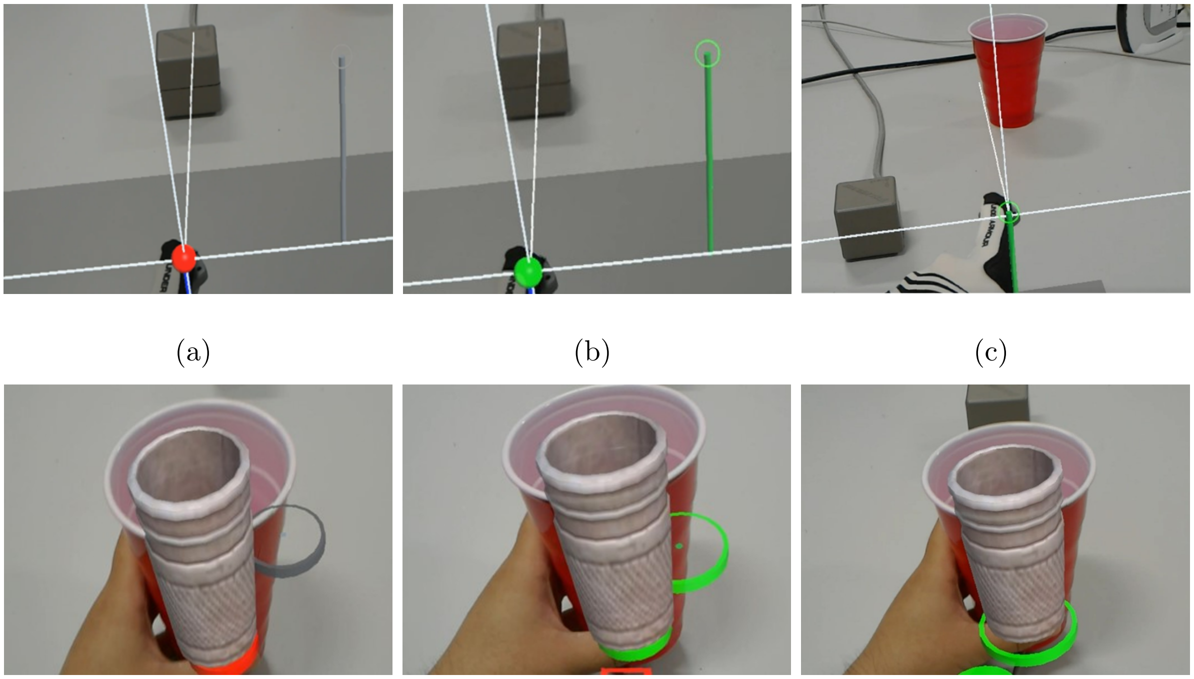
Fourteen able-bodied participants (five male and nine female, ages 19 - 32, one left-hand dominant, and thirteen right-hand dominant) provided written consent to participate in the study, which the Institutional Review Board approved. Each participant was asked to perform six different point-to-point reaching tasks: three separate conditions for the finger-pointing task, and three conditions for the cup-placement task.

2.2 Experiment Paradigm

We investigated two types of point-to-point reaching tasks: finger-pointing and cup-placement (Figures 1 and 2). For both tasks, three visualization conditions were tested:

(1) "MR with indicator" condition, in which participants perform point-to-point reaching tasks with a virtual fingertip or cup indicator displayed in an MR environment (Figure 1).

(2) "MR without indicator" condition that requires participants to perform point-to-point reaching tasks without any virtual indicators in an MR environment (Figure 2[a] and



(a) (b) (c)

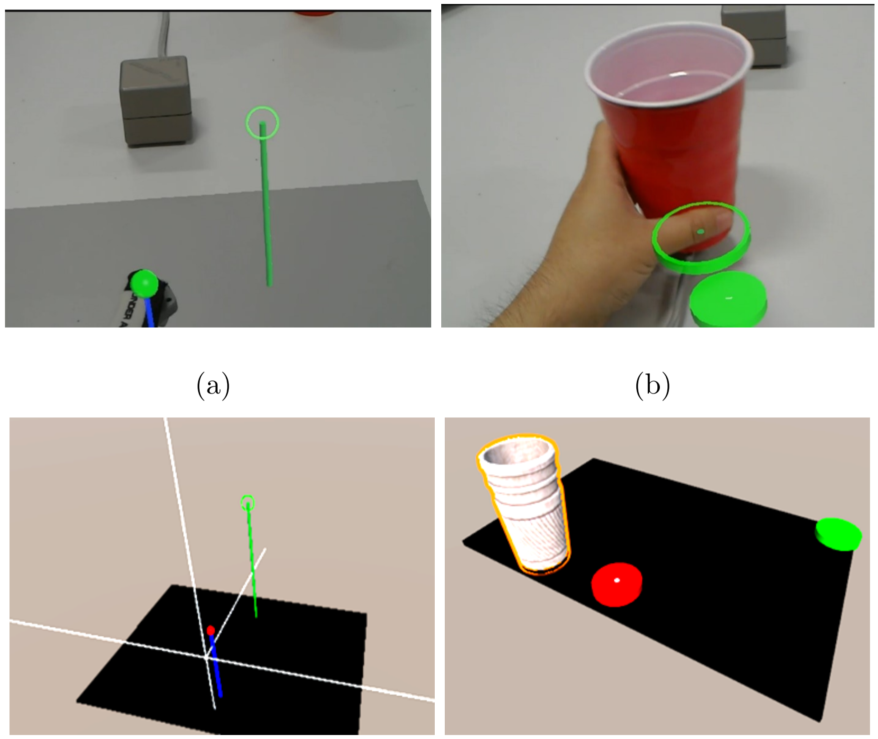
(d) (e) (f)

Figure 1: Participant's first-person view of the finger-pointing and cup-placement tasks. Sequences for the "MR with indicator" condition are showing (a and d) the fingertip/cup with indicator placed at the home position, which is shaded red and an inactive target that is shaded grey. (b and e) After participants click a remote control button, the home position and target indicators turn green and the participant will move to the target. (c and f) The participant reaches the target and clicks the remote button again.

[b]).

(3) "VR condition" that requires participants to perform the same tasks as in MR conditions with a virtual fingertip or cup indicator displayed, but in a VR environment where the veridical environment is not visible (simulations in Figure 2[c] and [d]).

For both tasks, participants sat in front of a veridical table and wore a head-mounted MR display (Microsoft HoloLens ver. 1, Microsoft Corp., Redmond WA), as seen in Figure



(a) (b)

(c) (d)

Figure 2: First person view of other experiment conditions for (a) finger-pointing MR without indicator (b) cup-placement MR without indicator tasks (c) finger-pointing VR with indicator (d) cup-placement VR with indicator tasks.

3. During the finger-pointing task, participants wore a glove instrumented with a position sensor (Liberty, Polhemus, Colchester, VT) on their dominant hand that tracks their index finger position. As for the cup-placement task, participants needed to hold a plastic cup that was instrumented with an identical position sensor attached under its base. Participants were provided with up to 5 minutes to practice making reaching movements before the actual experiments in both MR conditions for both finger-pointing and the cup-placement task. There was no practice for VR condition since it is very similar to MR condition with indicator. Participants were also asked to perform all the reaching tasks

as quickly and accurately as possible. For the finger-pointing task, each of the three experiment conditions included a set of 60 point-to-point reaching trials (60 targets). For each trial, participants were first presented with a red sphere at a home position that was consistent across all trials and a gray circle as the target (Figure 1[a]). Participants were instructed to put their index finger at the home position, then press a button on a remote control held by their non-dominant hand when they were ready to move to the target. Pressing the button changed the color of both the home position and target indicators to green and triggered audio feedback to confirm the button press (Figure 1[b]). Participants were instructed to press the button again when they believed that their fingertip reached the center of the target (Figure 1[c]). After this second button press, which triggers an audible tone distinct from the first press, the home position indicator's color becomes red (prompting participants to return to the home position) and a new target is presented at this location with a gray-colored indicator. The cup-placement task procedure was very similar to the finger-pointing task, but each visualization condition consisted of 20 point-to-point reaching trials (20 targets). Another difference was that the home position (a flat cylinder) and targets (a hollow cylinder) of the cup-placement task were both visualized to appear on the surface of a veridical table.

For both tasks, the home position was located near the edge of the veridical table and was vertically aligned with the glove's electromagnetic sensor's source. The experimental targets were constrained by a spherical sensing area with a radius of 65 cm. The 60 randomly ordered targets of the finger-pointing task were constructed from four repetitions of 15 unique targets, which had an index of difficulty range of 1.26–4.48 (Table 1), calculated with the modified Fitts' law discussed below. In order to minimize the effect of directional

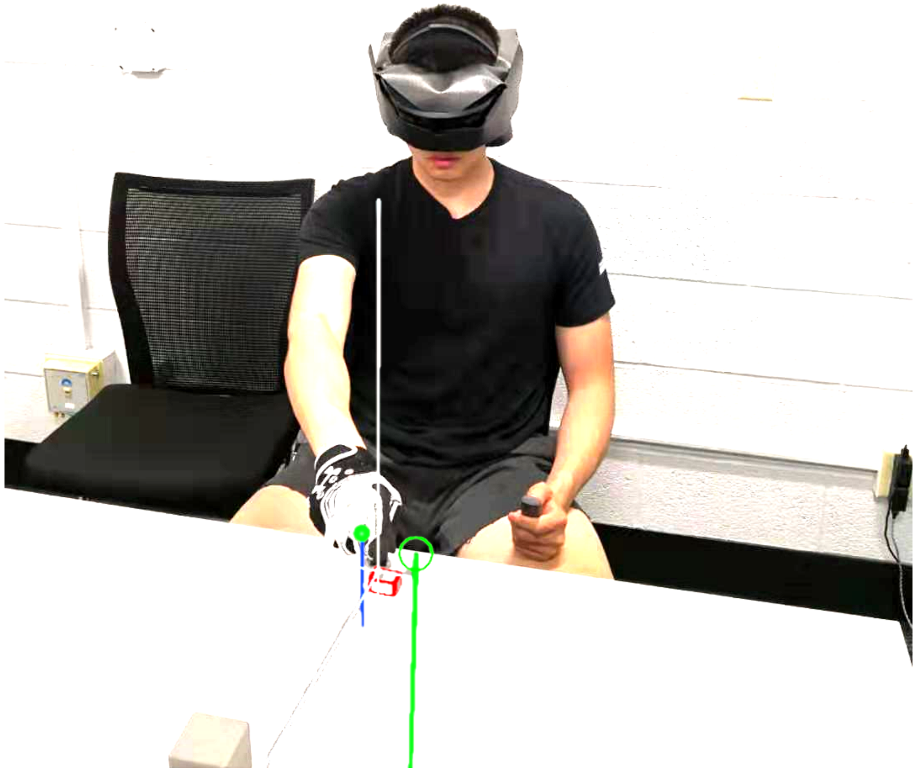


Figure 3: Example of the veridical and virtual work space for the finger-pointing task, with HoloLens shrouded for the VR modality. The virtual objects are illustrated as they appear to the participant.

bias, two repetitions of each target were mirrored and placed symmetrically on the opposite side of the home position. The 20 targets of the cup-placement task were constructed from one repetition of 10 unique targets that were mirrored to appear symmetrically on the opposite side of the home position, each with an index of difficulty range of 1-2.44 (Table 2).

Visualization indicators shown in Figure 2 were selected to be the simplest task representation while utilizing features to improve depth perception for the participant. Studies have shown that the design of mixed reality applications can affect the depth perception of veridical objects, which could confound results (Renner, Velichkovsky, & Helmert, 2013; Jones, Swan, Singh, Kolstad, & Ellis, 2008)

The cross indicator in the finger-pointing tasks ("MR with Indicator" and "VR" visualization conditions) was selected as it includes edges and vertices, which have been shown to improve depth perception (Ping, Weng, Liu, & Wang, 2020). Since the cup-placement task is representative of rehabilitation tasks, the cylinder indicator ("MR with Indicator" and "VR" visualization conditions) includes texture, curvature, and shading, traits of indicators that have been demonstrated to improve depth perception in mixed reality (Diaz, Walker, Szafir, & Szafir, 2017).

2.3 Reaching Task Targets Based on Fitts' Law

Fitts' index of difficulty was used to design the placement of the targets for this experiment's finger-pointing and cup-placement tasks (Tables 1 and 2). In the original Fitts' task, the participants were asked to use a stylus to make contact with a series of targets as quickly and accurately as possible (Fitts, 1954; Fitts & Peterson, 1964). In this study, we applied Shannon's formulation of Fitts' law, which describes reaching movement time MT to point at a target as:

$$MT = a + b * log_2 \left(1 + \frac{D}{W}\right),$$

where D is the distance between the home position and the center of the target, W is the width of the target, and the constants a and b are empirically determined by regression analysis. The term $log_2 \left(1 + \frac{D}{W}\right)$ is referred to as the index of difficulty (ID). ID rep-

Table 1: Finger-pointing task targets

ID (bits)	Distance (mm, home to center)	Diameter (mm)
1.26	42	30
1.49	54	30
1.79	69	30
1.95	72	25
2.18	88	25
2.41	108	25
2.64	105	20
2.87	126	20
3.10	152	20
3.33	136	15
3.56	162	15
3.79	193	15
4.02	152	10
4.25	180	10
4.48	213	10

Table 2: Cup-placement task targets

ID (bits)	Distance(mm)	Diameter (mm)
1.00	70	70
1.16	86.5	70
1.32	97.5	65
1.48	116.5	65
1.64	127	60
1.80	149.5	60
1.96	159.5	55
2.12	184.5	55
2.28	193	50
2.44	221	50

resents the difficulty level of the movement required by each target, and the units of ID are bits. To account for the observation that participants tend to end point-to-point movements closer to the edge of wider targets for point-to-point tasks (Soukoreff & MacKenzie, 2004), we computed the effective index of difficulty (ID_e) for each unique movement condition as:

$$ID_e = log_2\left(1 + \frac{D_e}{W_e}\right),$$

where D_e is the mean movement distance over all trials for each target movement condition and $W_e = \sigma\sqrt{2\pi e}$, where σ is the standard deviation of the endpoints. End-point distribution was assumed to be a normal random distribution. Fitts' law was originally applied to one dimension of movement, but has since been well-established to apply also to movement in two and three dimensions (MacKenzie, 1992; Murata & Iwase, 2001; M. Fu et al., 2012; Grossman & Balakrishnan, 2004; Cha & Myung, 2013) and to hold for movement while grasping an object (Smeets & Brenner, 1999; Thumser, Slifkin, Beckler, & Marasco, 2018).

2.4 Equipment

HoloLens was selected to visualize both MR and VR modalities in this study to minimize confounding effects related to optics, device mass, or field of view (34 W x 17 H degrees) across experiment conditions. No modifications were made to HoloLens for MR experiment conditions, but for VR experiment conditions we covered the device with a custom shroud that obstructed the participant's view of the veridical environment (Figure 3). The shroud was made of opaque 1/4 inch-thick foam, did not obstruct the device's cameras and sensors, and was adhered using Air Stick microsuction tape (Sewell Development Corp., Provo UT).

We used a USB Polhemus Liberty electromagnetic sensor to track finger or cup position and orientation because HoloLens Version 1 does not track hand position unless a hand gesture is being performed. While we could have used HoloLens Version 2 for optical hand tracking, this tracking fails if the fingers are obstructed by a cup when grasping or if they go outside the onboard camera's field of view. The sensor's magnetic source was placed in the center of a wood-surfaced table and the virtual work space was visually calibrated to appear on the table surface between the magnetic source and the participant before each participant's use (Figure 3). Velcro tape was used to attach a sensor to the index finger of a baseball batting glove (Under Armour Inc., Baltimore, MD) for finger-pointing trials and to the bottom of a plastic cup for cup-placement trials.

An Alienware 15 R3 Windows 10 Professional laptop (Dell Inc., Round Rock, TX) ran a custom Unity3D server application to sample the tracking sensor at 120 Hz and button presses from a Spotlight Bluetooth remote (Logitech, Newark, CA) - all of which was transferred to HoloLens via WiFi for visualization and logged locally.

2.5 Performance Measures

Throughput where x is the number of movement conditions (reaching targets), and y is the number of participants. $\mathrm{ID_e}$ is the effective index of difficulty computed from the empirical end-point distribution and participant reaching distance. TP in each task was computed for each participant per visualization condition for statistical analysis.

End-point error (EPE) is the Euclidean distance between the end point of a movement path and the target's central location. Lower EPE values were considered to reflect increased task performance.

Number of corrective movements (CMs) is defined as the number of direction changes (determined by the local maxima of the acceleration signal) of each movement trajectory. The smoother the movement, the lower the number of CMs. This analysis method was used in Blackmon, Cavusoglu, Fuji Lai, and Stark (1997); M. Fu et al. (2012) to measure human reaching performance in both virtual and real environments, with a larger number of CMs indicating worse task performance. In our case, we find the local maxima of acceleration (second derivative of position) using the "argrelextrema" method in the SciPy Python library. Velocity was determined based on the first derivative of position, which was then filtered by a 5 Hz third-order lowpass Butterworth filter using the signal.butter() and signal.filtfilt() functions from the Python SciPy library. Acceleration was determined based on the first derivative of velocity, which was then filtered with the same 5 Hz lowpass filter. To account for asynchrony in human movement, the final value was calculated by adding the number of corrective movements on each axis and dividing by three for the number of axes of motion. Since the CMs can be correlated with target difficulty (M. Fu et al., 2012), we also consider the possible effect of ID on the CMs.

Initial movement error (IME) is the magnitude of the difference between two normalized vectors: the participant's initial movement vector and the vector from the target vector. The initial movement vector points from the initial movement position to the first corrective movement position. The target vector points from the initial movement position to the current target position. Since increased IME may result in a longer reaching movement path, we considered greater IME to reflect reduced task performance.

Peak velocity (PV) is defined as the greatest absolute value of velocity for each reaching trajectory. Based on Blackmon et al. (1997); M. Fu et al. (2012), we consider higher PV to reflect better motor control confidence and increased task performance. To calculate the PV, we take the squared root of the sum of squares of the peak velocity on each axis.

Efficiency is defined as how far the actual movement path deviated from the direct path to the target. The first formula of efficiency was defined in Zhai and Milgram (1998). In our case, the formula is

$$\text{Efficiency} = \frac{D_{\text{endpoint}}}{D_{\text{path}} - D_{\text{endpoint}}}, \tag{1}$$

where $D_{endpoint}$ is the Euclidean distance from the initial movement point to the end-point of the reaching motion. D_{path} indicates the length of the path between the initial and end-points. Higher efficiency means the movement path is more likely to be a straight line, so we assumed that increased efficiency means better task performance.

2.6 Data Pre-Processing

Prior to analysis, trajectories from each trial were pre-processed to eliminate dwelling behavior caused by participants keeping their finger or cup on the start or target positions prior to moving. To remove dwell time, only trajectories with a velocity higher than a threshold of 1.5mm/s were used in the analysis. This 1.5mm/s velocity threshold was chosen based on expert retinal surgeons' hand tremor frequency when they hold a stylus grip (Blackmon et al., 1997). Also, eight outlying trials were removed because the Bluetooth remote control button press meant to indicate the end of a trial failed to be registered by the server. In these trials, trajectories included not only the path from the home position to the target but also the extraneous return path from the target back to the home position.

2.7 Statistical Analysis

The data were determined to be non-normally distributed via the Shapiro-Wilk test, hypothesis tests were performed for each performance measure to compare mean differences between each visualization condition using Friedman Tests and Wilcoxon rank sum post-hoc tests. All tests were performed using R 4.3.0 with the rstatix, lmerTest, and robustlmm libraries. Statistical power for each performance measure was computed to be 0.52 (Statistical power was calculated by G*Power 3.1 (Faul, Erdfelder, Lang, & Buchner, 2007) with Cohen's medium effect size of 0.25, $\alpha = 0.05$, sample size of 15, one group, and three repeated measurements).

Due to the replicated nature of the data, the means of each outcome measure were calculated for each ID, creating an unreplicated block design for the non-parametric Friedman Test evaluation of statistical significance. For each performance measure, linear regression was performed to find the relationship between ID and the performance measure for each experiment condition. Slope was interpreted as the sensitivity of the performance measure to the range of tested IDs. Offset was interpreted as the performance value for the median tested ID and was computed from the regression slope, m, and regression zero intercept, $y_\circ$, as

$$\text{offset} = y_\circ + m\left(\text{ID}_{\text{min}}\right), \tag{2}$$

where ID_{min} was the median target ID for each task, with 2.87 bits for finger-pointing, Table 1, and 1.72 bits for cup-placements Table 2. Offset was conceived by Guiard and Olafsdottir (2011) to link regression zero intercept to empirical results because a target with zero ID cannot be constructed.

If target ID had a significant effect on a performance measure, then the Friedman test was performed with the linear regression slope as the dependent variable and experiment condition as the within-subjects factor. If target ID did not have a significant effect on the performance measure, then the Friedman test was performed with the performance measure as the dependent variable and experiment condition as the within-subjects factor.

3 Results

3.1 Throughput

Target difficulty (ID) was found to have a significant effect on movement time for both finger-pointing and cup-placement tasks F(14, 45) = 40.433, ($p = 0.001$, and F(9, 30) = 21.182, $p = 0.012$. Therefore, robust linear regression was performed and statistical tests for significant differences were calculated using throughput as the dependent variable and experiment conditions as the within-subjects factor.

Finger-pointing task As shown in Figure 4[a] and Table 3, statistically significant differences in throughput were not found between visualization conditions, F(2, 42) = 1.714, $p = 0.607$. Although not statistically significant, the mean TP was highest for the "MR without indicator" condition (1.11 bits/s), followed by the "VR" (1.01 bits/s) and "MR with indicator" (0.92 bits/s) conditions.

Cup-placement task As shown in Figure 5[a] and Table 4, statistically significant differences in throughput were not found between visualization conditions F(2, 42) = 1, $p = 0.4244$. Although not statistically significant, the mean TP for both "MR with indicator" (0.95 bits/s) and "MR without indicator" (0.93 bits/s) conditions was higher than the "VR" condition (0.85 bits/s).

Table 3: Means ± Standard Deviations of Finger Pointing Tasks' Performance Measures

	MR with Indicator	MR without Indicator	VR
TP (bits/s)	0.917 ± 0.159	1.110 ± 0.150	1.008 ± 0.159
EPE Slope (m/bit)	0.003 ± 0.001	0.002 ± 0.000	0.002 ± 0.001
EPE offset (m)	0.017 ± 0.002	0.023 ± 0.001	0.021 ± 0.002
IME Slope (m)	0.026 ± 0.005	0.017 ± 0.006	0.016 ± 0.005
IME Offset (m)	0.043 ± 0.015	0.075 ± 0.014	0.063 ± 0.015
# of CM Slope	0.000 ± 0.000	0.000 ± 0.000	0.002 ± 0.000
# of CM Offset	1.000 ± 0.000	1.000 ± 0.000	1.000 ± 0.000
PV (m/s)	4.525 ± 0.008	4.633 ± 0.006	4.564 ± 0.008
Efficiency Slope	0.031 ± 0.014	0.072 ± 0.014	0.048 ± 0.014
Efficiency Offset	0.784 ± 0.104	0.682 ± 0.096	0.786 ± 0.104

Table 4: Means ± Standard Deviations of Cup Placement Tasks' Performance Measures

	MR with Indicator	MR without Indicator	VR
TP (bits/s)	0.950 ± 0.132	0.931 ± 0.111	0.852 ± 0.132
EPE (m)	0.010 ± 0.000	0.011 ± 0.000	0.012 ± 0.000
IME (m)	0.201 ± 0.012	0.122± 0.009	0.199 ± 0.012
# of CM	1.292 ± 0.048	1.197 ± 0.038	1.389 ± 0.048
PV (m/s)	1.709 ± 0.016	1.714 ± 0.011	1.787 ± 0.016
Efficiency Slope	0.233 ± 0.157	0.429 ± 0.111	0.311 ± 0.157
Efficiency Offset	0.851 ± 0.185	0.807 ± 0.186	0.826 ± 0.185

3.2 End-point Error

Target ID was shown to have a statistically detectable effect on end-point error for the finger-pointing task but not the cup-placement task, F(14, 45) = 33.867, $p \leq 0.001$ and F(9, 30) = 16.164, $p = 0.0635$.

Finger-pointing task visualization conditions did not have a statistically significant effect on the end-point error regression slope, but there was a trend that the "MR with indicator condition" (0.003) had higher mean slope than the other two conditions (0.002).

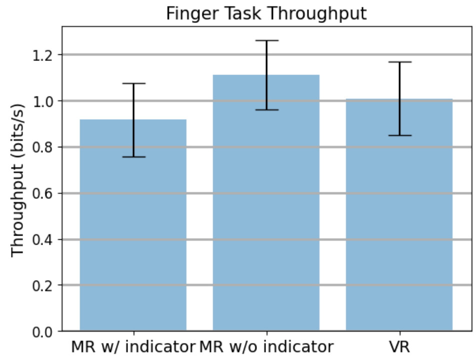
Visualization conditions had a significant effect on the end-point error linear regression offset, F(2, 42) = 8.7636, (p = 0.0125. Post-hoc analysis with Bonferroni correction shows that the "MR without Indicator" condition had a significantly higher regression offset (0.023±0.001) than the "MR with Indicator" condition (0.017 ± 0.002) and the "VR" condition was not statistically separable from the other two experiment conditions (Figure 4[b] and Table 3).

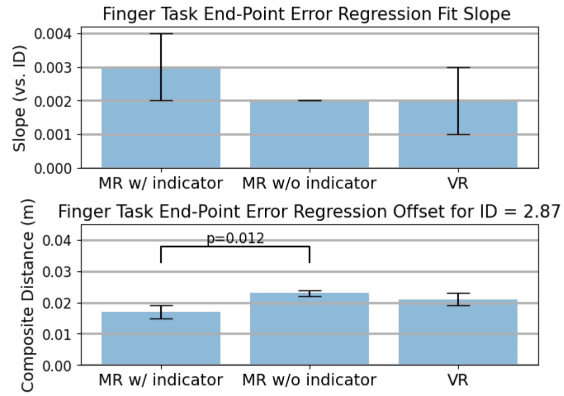
Cup-placement task visualization conditions had a significant effect on end-point error, F(2, 42) = 11.4 p = 0.003. The "MR without indicator" condition mean end-point error of 0.011 ± 0.000 was significantly higher than the "MR with indicator" (0.010 ± 0.000) conditions and significantly lower than the "VR" condition (0.012 ± 0.000) (Figure 5[b] and Table 4).

3.3 Initial Movement Error

Target ID had a significant effect on initial movement error for the finger-pointing task but not the cup-placement task, F(14, 45) = 34.567, p = 0.017 and F(9, 30) = 9.909, p = 0.358,.

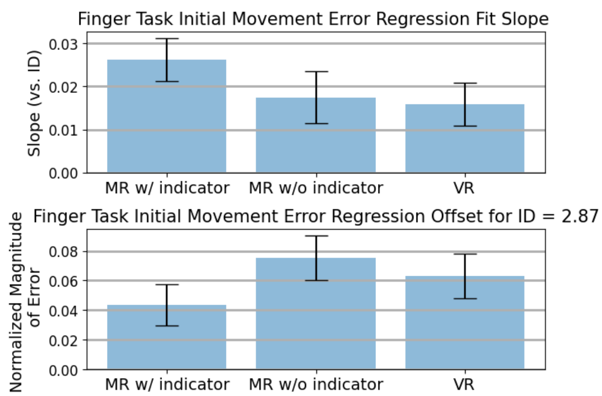
Finger-pointing task visualization conditions did not have a significant effect on either the regression slope or regression offset of initial movement error, F(2,42) = 3.444, p = 0.179. However, we can observe that the "MR with Indicator" condition was the most sensitive to ID (mean 0.026 slope), followed by the "MR without Indicator" condition (0.017 ± 0.006) and "VR" condition (0.016 ± 0.005) and had the lowest regression offset (0.043 ± 0.015). We also observed that the "MR without Indicator" condition had the

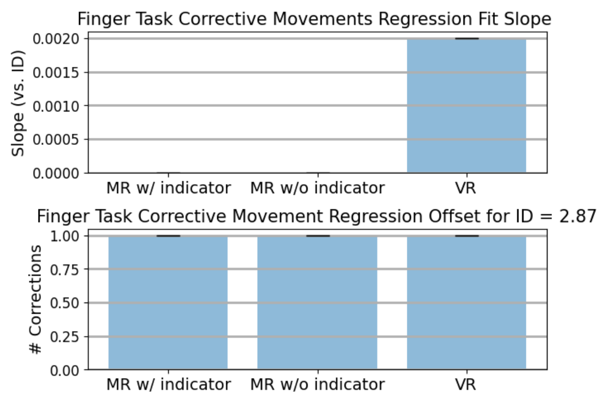




(a)

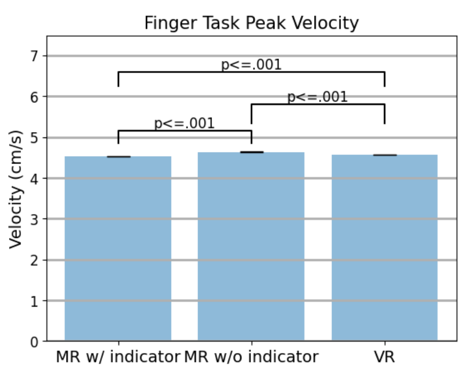
(b)

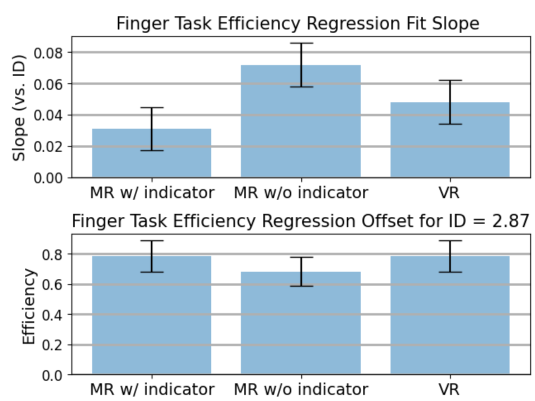




(c)

(d)





(e)

(f)

Figure 4: Performance measures for the finger-pointing task across all 3 experiment conditions. "I"-shaped whiskers indicate the standard deviations (±) from the means. Statistically significant multiple comparisons and their p-values are marked above the bars.

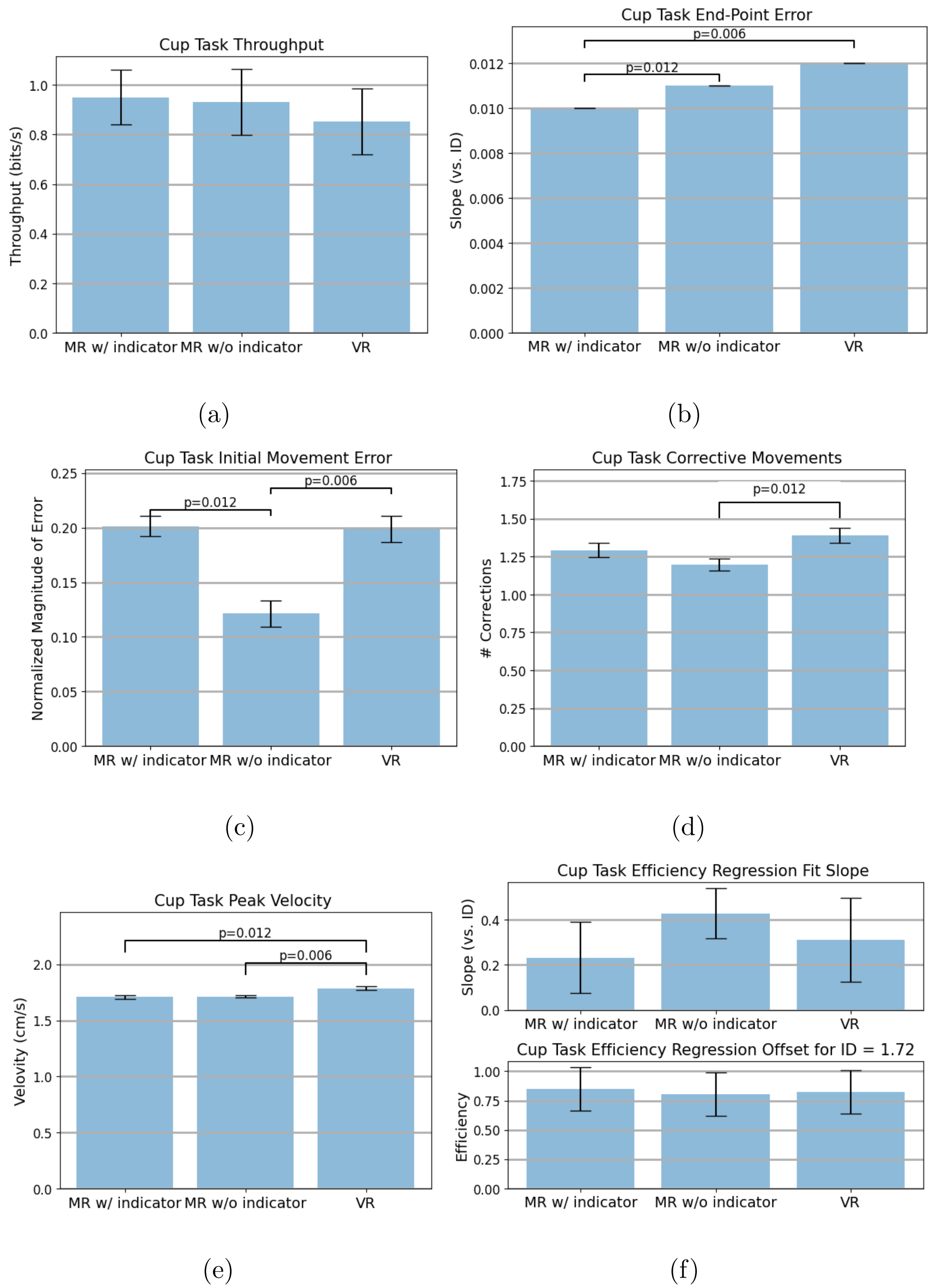


Figure 5: Performance measures for the cup-placement task across all 3 visualization conditions. "I"-shaped whiskers indicate the standard deviations (±) from the means. Statistically significant multiple comparisons and their p-values are marked above the bars.

highest regression offset (0.075 ± 0.014) compared to the other conditions. The "MR with Indicator" condition had the lowest regression offset (0.043 ± 0.015) and the "VR" condition had a regression offset (0.063 ± 0.051)

Cup-placement task visualization conditions did affect the initial movement error, F(2, 42) = 12.2, $p = 0.002$. Post-hoc analysis with Bonferroni correction revealed that the "MR without Indicator" condition had statistically lower initial movement error (0.122 ± 0.009) than the "MR with Indicator" (0.201 ± 0.012) and "VR" (0.199 ± 0.012) conditions. Further analysis of the trajectories indicated that participants tended to move the cup in an arc to the targets (rather than slide the cup toward the target) even though the targets were all located on the table surface (Figure 6). In this case, larger initial movement error can be interpreted as greater arcing in the participant's movement trajectory.

3.4 Corrective Movements

Target ID had a significant effect on the number of corrective movements for the finger-pointing task but not the cup-placement task, F(14,45) = 31.933, $p \leq 0.004$ and F(9,30) = 12.629, $p = 0.180$.

Finger-pointing task While experiment conditions were not found to have a significant effect on the regression slope and offset, it was observed that the "VR" condition (1.00 ± 0.00) had the highest regression slope, while the other visualization conditions were not affected by ID. The "MR with indicator" condition, and the "MR without Indicator" condition both had a regression slope of 0.00 ± 0.00. There was no difference across visualization conditions for regression fit offset, with each condition having a regression

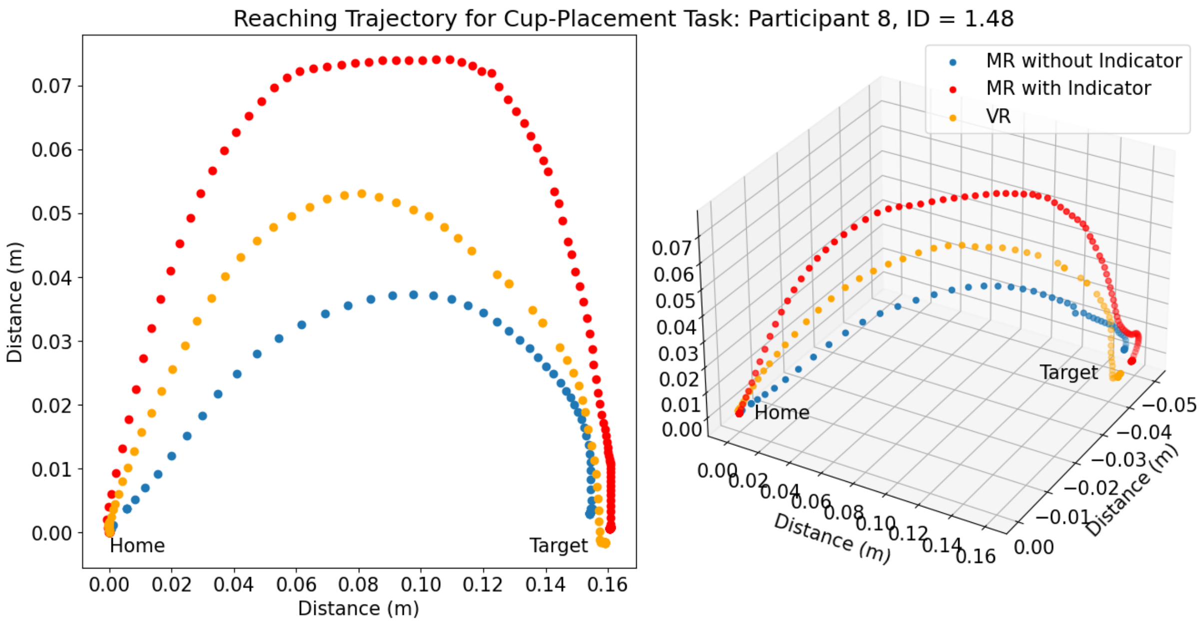


Figure 6: An example set of a participant's movement path for a cup-placement task to the same difficulty target (ID = 1.48) for different visualization conditions in the vertical and forward directions (left) and vertical, forward, and horizontal directions (right).

offset of 1.00 ± 0.00 as shown in (Figure 4[d] and Table 3).

Cup-placement task visualization conditions significantly impacted the number of corrective movements F(2, 42) = 12.629, p 0.018. Post-hoc analysis demonstrated that the "MR without Indicator" condition had statistically fewer corrective movements (1.197 ± 0.038) compared to the "VR" condition (1.389 ± 0.048). The "MR with Indicator" condition (1.292 ± 0.048) was not statistically separable from the other conditions.

3.5 Peak Velocity

Target ID did not have a significant effect on peak velocity for the finger-pointing or cup-placement tasks, F(14,45) = 11.667, $p = 0.633$ and F(9,30) = 8.382), $p = 0.496$.

Finger-pointing task experiment conditions were found to have significant effects on peak velocity F(2, 42) = 28.133, $p \leq 0.001$. Multiple comparisons showed significant differences between each experiment condition. The "MR without Indicator" had the highest peak velocity (4.633 $\pm$ 0.006) m/s, followed by the "VR" condition (4.564$\pm$ 0.008) m/s, while the "MR with indicator" had the lowest peak velocity (4.525 $\pm$ 0.008) m/s. Post-hoc analysis revealed that all three conditions were statistically separable (Figure 4[e] and Table 3).

Cup-placement task experiment conditions significantly impacted peak velocity, F(2, 42) = 12.2, $p = 0.002$. Post-hoc analysis demonstrated that the "VR" condition had a statistically higher peak velocity (1.787 $\pm$ 0.016) m/s than the "MR with Indicator" (1.709 $\pm$ 0.016) m/s and the "MR without Indicator" condition (1.714 $\pm$ 0.011) m/s (Figure 5[e] and Table 4).

3.6 Efficiency

Target ID had a significant effect on efficiency for the finger-pointing and cup-placement tasks, F(14, 45) = 39.1, $p = 0.002$ and F(9, 30) = 26.005), $p = 0.23$.

Finger-pointing task experiment conditions did not have a statistically separable effect on the efficiency regression slope. When observing the regression fit slope for each condition, the "MR with indicator" condition had the lowest regression slope (0.031 $\pm$ 0.014), and the "MR without Indicator" condition had the steepest slope (0.072 $\pm$ 0.014). The regression slope for the "VR" condition was 0.048 $\pm$ 0.014. Additionally, experiment conditions, while not statistically separable, demonstrated that the "VR" condition showed the highest efficiency regression offset (0.786 $\pm$ 0.104), followed by the "MR with indicator" condition (0.784 $\pm$ 0.104), and lastly the "MR without Indicator" (0.682 $\pm$ 0.096) had the lowest regression offset. The effect of the (Figure 4[f] and Table 3).

Cup-placement task experiment conditions, while not statistically separable, demonstrated that the "MR without indicator" condition showed the highest efficiency regression slope (0.429 $\pm$ 0.111), followed by the "VR" condition (0.311 $\pm$ 0.157), and lastly the "MR with Indicator" condition (0.233 $\pm$ 0.157). The effect of the experiment condition on the efficiency regression offset was also not statistically separable. However, unlike the finger-pointing task, the "MR with Indicator" condition had the highest regression offset (0.851 $\pm$ 0.185), and the "MR without indicator" condition had the lowest regression offset (0.807 $\pm$ 0.186). The "VR" experiment condition had a regression offset of 0.826 $\pm$ 0.185 (Figure 5[f] and Table 4).

4 Discussion

Limitations of the study include limited sample size ($n = 14$) and the use of a pseudo-VR visualization condition. A more effective experimental design may have used a pass-through VR mode that shows the participant's surroundings through a video stream, which may limit confounding depth perception effects of objects in the environment.

The results provided unexpected evidence against our hypothesis, which we discuss in more detail below.

4.1 Finger-Pointing Task

Compared with M. Fu et al. (2012)'s co-located VR study condition, our "VR" condition's results had increases in throughput, peak velocity, and corrective movement offset; along with decreases in efficiency and corrective movement slope. The difference in performance measures could be due to differences in the experiment paradigm, as M. Fu et al. (2012) used a stylus for pointing and the co-located VR display was a fixed stereoscopic computer monitor reflected through a mirror.

While the regression fit offset demonstrated a lower end-point error when the target ID was 2.87 bits, the mean regression fit slope for end-point error for the "MR with indicator" condition was higher than the "MR without indicator"and "VR" conditions, indicating a higher sensitivity with the inclusion of an indicator.

Compared to related Fitts' task literature, we found the performance of our "MR with indicator" condition had a slight increase in throughput and two times more decrease in end-point error compared with Ha and Woo (2010)'s best virtual hand avatar visualization techniques. This discrepancy may be because reaching performance under a co-located experiment condition is better than under non-colocated (M. Fu et al., 2012). Alternatively, this may indicate a lack of synchronization between the indicator and fingertip at larger distances.

As for initial movement error, the "MR without indicator" condition was higher than the "MR with indicator" and the "VR" conditions. While not statistically significant, the inclusion of the fingertip indicator in the MR environment decreased efficiency performance for the regression fit offset, but increased sensitivity when compared to the "MR without Indicator" and "VR" conditions.

Corrective movement values remained consistent across display modalities, however the "VR" visualization condition was the most sensitive.

The results revealed a trend that the fastest finger-pointing trajectories occurred when the fingertip indicator was not provided in the MR environment. Specifically, peak velocity for the "MR without indicator" condition was statistically higher than the other two conditions. The "MR without indicator" condition also exhibited trends of faster and smoother motion than the other two conditions, which was indicated by the highest mean throughput and lowest mean number of corrective movements regression slopes and offsets, respectively than the other two conditions.

Finally, we observed that the "MR without Indicator" visualization condition was more sensitive to ID than the "VR" and "MR without Indicator" conditions when considering efficiency performance. This indicates that participants took a more direct path when the visualization condition did not include an indicator.

4.2 Cup-Placement Task

Many results for the cup-placement task did not reach statistical significance, but some interesting trends were observed.

The mixed reality experiment conditions had higher throughput than the virtual reality condition. While not statistically separable, the "MR with Indicator" condition had the highest throughput, followed by the "MR without Indicator" condition.

We also observed that participants had the most difficulty reaching the target location in an MR environment when not presented with an indicator. From end-point error results, we found that the mean error for "MR without indicator" condition was higher than the "MR with indicator" condition and "VR" condition. This trend was also similar to finger-pointing tasks.

For initial movement error, we observed that the mean initial movement errors for the "MR without Indicator" condition were statistically lower than the "MR with Indicator" and "VR" conditions. When comparing to the finger-pointing task, the higher initial movement error appears to be due to an arcing trajectory that had the highest amplitude for the "MR with indicator" condition, followed by the "VR" condition, and then the "MR without indicator" condition for the same target ID (illustrated in Figure 6), which can lead to large initial movement errors in vertical movement relative to forward and lateral movement, and potentially confound differences between visualization conditions.

Similar to the finger-pointing task, the "MR without Indicator" condition had significantly fewer corrective movements when compared to the "VR" condition. While not significant, we can observe that participants under the "MR without indicator" condition exhibited smoother reaching motion than the "MR with indicator" and "VR" conditions for the cup-placement task.

Additionally, the "VR" condition had the fastest peak velocity of the three conditions, which is in agreement with the results from the finger-pointing task. Finally, similar to the finger-pointing tasks, the "MR without Indicator" visualization was most sensitive to ID for efficiency.

4.3 Trade-off between Reaching Confidence and Accuracy

These results suggest a trade-off between reaching motion confidence and accuracy that is well documented in human movement behavior in veridical and virtual environments for reaching and object manipulation (Plamondon & Alimi, 1997; Chang, Hsu, Hsu, & Chen, 2016) and further research has suggested that Fitts' throughput may be independent of this speed-accuracy trade-off (MacKenzie & Isokoski, 2008). This trade-off refers to the possibility that participants prioritized movement speed at the expense of accuracy when the fingertip indicator was not visible and, vice-versa, prioritized accuracy at the cost of speed in the experiment when the fingertip indicator was visible. Additional studies are needed to explore trade-off across visualization conditions fully, but initial support from the finger-pointing task comes from noticing that throughput, peak velocity, and efficiency (slope) were best in the "MR without indicator" condition and can be considered as indicators of what we term "movement confidence". Similarly, in the cup placement task, the "MR without Indicator" condition had the best performance for the efficiency (slope) and corrective movements measures. Additionally, the "MR without Indicator" also performed poorly when looking at measures of movement accuracy (endpoint error and initial movement error offsets) for the finger-pointing task. This trend was not observed for the cup-placement task.

This did not support our hypothesis that displaying the indicator in the "MR with indicator" condition would result in the highest reaching performance on all measures. However, this unexpected result may indicate that providing participants with an endpoint indicator causes them to prioritize accuracy over movement speed. On the other hand, participants under "MR without indicator" prioritized moving fast at the cost of accuracy. Although we observed the trend that participants under the "MR without indicator" condition had faster and smoother reaching motions and reduced accuracy, additional studies are required to validate this trend.

The implications of this finding on MR/VR application design are important. When both MR and VR are under consideration, MR is suitable if precise end-point accuracy is not required, such as the training of fast, ballistic movements – as required in gross motor tasks or rehabilitation movement training. If the features of MR are necessary, then displaying an end-point indicator is recommended for maximum performance. On the other hand, if MR is not necessary and end-point error is critical, such as fine motor tasks or precise rehabilitation movement training, then a VR-based environment design may be most appropriate.

4.4 Relevance to Rehabilitation

These findings can be translated to MR- and VR-based rehabilitation applications. Common occupational therapy exercises are designed to increase dexterity, range of motion, and strength. The purpose of MR- and VR-based rehabilitation is to augment the experience with feedback to the user in a way that is fun and interactive, promoting extended engagement and repetition (Lang, 2009). In stroke rehabilitation, common rehabilitation measures such as the Box and Block Test and the Nine-Hole Peg Test assess dexterity and upper extremity function by counting the number of tasks completed in a set period. These are tests that assess both speed and accuracy and highlight the importance of both measures when it comes to physical rehabilitation. The findings from this study can be applied to VR- and MR-based adaptations of rehabilitation exercises.

Using MR-based display modalities to perform exercises such as picking up small objects, such as coins or paperclips, to include targets, point systems, and other gamification techniques rehabilitation specialists can utilize MR for tasks that require high levels of accuracy. Other rehabilitation activities that might leverage MR to perform tasks with weighted objects, such as pulling a heavy object across a table or raising objects with both hands may not require accuracy and instead focus on mass repetition.

Rehabilitation exercises such as holding bottles or hammer-like objects, or lifting objects with both hands, can be informed by results from the cup-placement tasks. Since these applications are most similar to the cup-placement task, these applications should utilize findings from that task. Specifically, if accuracy is preferred and participants are intended to minimize end-point error and employ many corrective movements, the VR-based display modality should be used. Additionally, if speed and repetition are prioritized and participants are intended to maximize velocity without regard to accuracy, the study design should use a VR-based display modality.

Exercises to improve range of motion often include multiple repetitions of movements, such as moving the hand to the stomach, mouth, or table. These point-to-point reaching tasks can be informed by results from the finger-pointing tasks, and often prioritize speed. Therefore, tasks should be presented in a mixed reality context without indicator. However, if accuracy and fine motor control are preferred, tasks should be presented to increase corrective movements and decrease initial movement and end point error. Therefore, these tasks should be presented in virtual reality.

5 Conclusion

This study examined six different performance measures for two types of point-to-point reaching tasks (finger-pointing and cup-placement tasks) with three environment conditions based on Fitts' law. A key finding of this study was evidence of a potential trade-off between reaching motion confidence and accuracy. Specifically, participants tended to move fastest and smoothest in the MR without finger or cup indicator conditions, but their endpoint error was also highest. Additionally, in conditions where an indicator was provided in MR, finger-pointing and cup-placement task performance was comparable to the VR condition. These results are important because they suggest that MR tasks requiring low end-point error require an indicator to be provided in the user interface. However, if end-point accuracy is not critical to the application, then not visualizing the endpoint can improve users' movement speed and smoothness.

6 Acknowledgements

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. 1937968, CBET 1942402 and the MetroHealth Foundation Neurorehabilitation Research Fund. We extend our gratitude to the Case Western Reserve University Interactive Commons for providing access to the Microsoft HoloLens.

References

Blackmon, T. T., Cavusoglu, M. C., Fuji Lai, & Stark, L. W. (1997). Human hand trajectory analysis in point-and-direct telerobotics. In *1997 8th International Conference on Advanced Robotics Proceedings ICAR'97* (p. 927-932).

Cha, Y., & Myung, R. (2013, 07). Extended fitts' law for 3d pointing tasks using 3d target arrangements. *International Journal of Industrial Ergonomics*, *43*, 350–355. doi: 10.1016/j.ergon.2013.05.005

Chang, C.-M., Hsu, C.-H., Hsu, C.-F., & Chen, K.-T. (2016). Performance measurements of virtual reality systems: Quantifying the timing and positioning accuracy. In *Proceedings of the 24th ACM International Conference on Multimedia* (p. 655–659). New York, NY, USA: Association for Computing Machinery. Retrieved from `https://doi.org/10.1145/2964284.2967303` doi: 10.1145/2964284.2967303

Diaz, C., Walker, M., Szafir, D. A., & Szafir, D. (2017). Designing for depth perceptions in augmented reality. In *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (p. 111-122). doi: 10.1109/ISMAR.2017.28

Duff, M., Chen, Y., Attygalle, S., Herman, J., Sundaram, H., Qian, G., … Rikakis, T. (2010, Oct). An adaptive mixed reality training system for stroke rehabilitation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *18*(5), 531-541. doi: 10.1109/TNSRE.2010.2055061

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *47*(6), 381. doi: 10.1037/h0055392

Fitts, P. M., & Peterson, J. R. (1964). Information capacity of discrete motor responses. *Journal of Experimental Psychology*, *67*(2), 103. doi: 10.1037/h0045689

Fu, M., Hershberger, A., Sano, K., & Çavuşoğlu, M. (2012, 08). Effect of visuo-motor co-location on 3d fitts' task performance in physical and virtual environments. *Teleoperators and Virtual Environments - Presence*, *21*, 305-320. doi: 10.1162/PRES$_{a0}$0115

Fu, M. J., Knutson, J. S., & Chae, J. (2015). Stroke rehabilitation using virtual environments. *Physical Medicine and Rehabilitation Clinics*, *26*(4), 747–757.

Grossman, T., & Balakrishnan, R. (2004, 01). Pointing at trivariate targets in 3d environments. *Conference on Human Factors in Computing Systems - Proceedings*, 447-454. doi: 10.1145/985692.985749

Guiard, Y., & Olafsdottir, H. B. (2011). On the measurement of movement difficulty in the standard approach to fitts' law. *PLoS One*, *6*(10), e24389.

Ha, T., & Woo, W. (2010). An empirical evaluation of virtual hand techniques for 3d object manipulation in a tangible augmented reality environment. In *2010 IEEE Symposium on 3D User Interfaces (3DUI)* (pp. 91–98).

Holden, M. K. (2005). Virtual environments for motor rehabilitation: Review. *Cyberpsychology Behavior and Social Networking*, *8*(3), 187-211. doi: 10.1089/cpb.2005.8.187

Jones, J. A., Swan, J. E., Singh, G., Kolstad, E., & Ellis, S. R. (2008). The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception. In *Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization* (p. 9–14). New York, NY, USA: Association for Computing Machinery. Retrieved from `https://doi.org/10.1145/1394281.1394283` doi:

10.1145/1394281.1394283

Khademi, M., Mousavi Hondori, H., Dodakian, L., Cramer, S., & Lopes, C. (2013, 07). Comparing 'pick and place' task in spatial augmented reality versus non-immersive virtual reality for rehabilitation setting. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, *2013*, 4613-4616. doi: 10.1109/EMBC.2013.6610575

Lang, R. B. J. K. S.-I. H. R. S., MacDonald. (2009). Observation of amounts of movement practice provided during stroke rehabilitation. *Archives of Physical Medicine and Rehabilitation*, *90*(10), 1692-1698.

MacKenzie, I. S. (1992, March). Fitts' law as a research and design tool in human-computer interaction. *Human-Computer Interaction*, *7*(1), 91–139.

MacKenzie, I. S., & Isokoski, P. (2008). Fitts' throughput and the speed-accuracy tradeoff. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (p. 1633–1636). New York, NY, USA: Association for Computing Machinery. Retrieved from `https://doi.org/10.1145/1357054.1357308` doi: 10.1145/1357054.1357308

Microsoft Corp. (2019). *Microsoft Mixed Reality getting around hololens (1st gen).* Retrieved 2021-01-22, from `https://docs.microsoft.com/en-us/hololens/hololens1-basic-usage`

Murata, A., & Iwase, H. (2001). Extending fitts' law to a three-dimensional pointing task. *Human Movement Science*, *20*(6), 791 - 805. doi: doi.org/10.1016/S0167-9457(01)00058-6

Ping, J., Weng, D., Liu, Y., & Wang, Y. (2020). Depth perception in shuffleboard: Depth cues effect on depth perception in virtual and augmented reality system. *Journal of the Society for Information Display*, *28*(2), 164-176. Retrieved from `https://sid.onlinelibrary.wiley.com/doi/abs/10.1002/jsid.840` doi: https://doi.org/10.1002/jsid.840

Plamondon, R., & Alimi, A. M. (1997). Speed/accuracy trade-offs in target-directed movements. *Behavioral and Brain Sciences*, *20*(2), 279–303. doi: 10.1017/S0140525X97001441

Pridmore, T., Cobb, S., Hilton, D., Green, J., & Eastgate, R. (2007). Mixed reality environments in stroke rehabilitation: Interfaces across the real/virtual divide. *International Journal on Disability and Human Development*, *6*(1), 87–96. Retrieved 2024-03-11, from `https://doi.org/10.1515/IJDHD.2007.6.1.87` doi: doi:10.1515/IJDHD.2007.6.1.87

Renner, R. S., Velichkovsky, B. M., & Helmert, J. R. (2013, dec). The perception of egocentric distances in virtual environments - a review. *ACM Computing Surveys*, *46*(2). Retrieved from `https://doi.org/10.1145/2543581.2543590` doi: 10.1145/2543581.2543590

Scalona, E., Hayes, D., Del Prete, Z., Palermo, E., & Rossi, S. (2019). Perturbed point-to-point reaching tasks in a 3d environment using a portable haptic device. *Electronics*, *8*(1), 32.

Smeets, J., & Brenner, E. (1999, 08). A new view on grasping. *Motor control*, *3*, 237-71. doi: 10.1123/mcj.3.3.237

Soukoreff, R., & MacKenzie, I. (2004, 12). Towards a standard for pointing device evaluation, perspectives on 27 years of fitts' law research in hci. *International Journal of Human-Computer Studies*, *61*, 751-789. doi: 10.1016/j.ijhcs.2004.09.001

Tang, X., Hu, X., Fu, C.-W., & Cohen-Or, D. (2020). Grabar: Occlusion-aware grabbing virtual objects in ar. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (p. 697–708). New York, NY, USA: Association for Computing Machinery. Retrieved from `https://doi.org/10.1145/3379337.3415835` doi: 10.1145/3379337.3415835

Thumser, Z. C., Slifkin, A. B., Beckler, D. T., & Marasco, P. D. (2018). Fitts' law in the control of isometric grip force with naturalistic targets. *Frontiers in Psychology*, *9*. Retrieved from `https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2018.00560` doi: 10.3389/fpsyg.2018.00560

Zhai, S., & Milgram, P. (1998). Quantifying coordination in multiple dof movement and its application to evaluating 6 dof input devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (p. 320–327). USA: ACM Press/Addison-Wesley Publishing Co. Retrieved from `https://doi.org/10.1145/274644.274689` doi: 10.1145/274644.274689

List of Tables

List of Figures