Hitting the Wall: Mid-Air Interaction for Eye-Hand Coordination

Anil Ufuk Batmaz  
abatmaz@sfu.ca  
Simon Fraser Univ., SIAT  
Vancouver, BC, Canada

Xintian Sun  
xintain_sun@sfu.ca  
Simon Fraser Univ., SIAT  
Vancouver, BC, Canada

Dogu Taskiran  
dogu@stambol.com  
Stambol Studios Inc.  
Vancouver, BC, Canada

Wolfgang Stuerzlinger  
w.s@sfu.ca  
Simon Fraser Univ., SIAT  
Vancouver, BC, Canada

ABSTRACT

Reaction time training systems are used to improve user performance. Until now, such setups use physical 2D flat surfaces, e.g., a 2D touch screen or buttons mounted on a wall. We designed and investigated a mid-air reaction time training system with an immersive virtual reality (VR) headset. 12 participants performed an eye-hand coordination reaction test in three conditions: both in mid-air with or without VR controller as well as with passive haptic feedback through hitting a soft-surface wall. We also altered target and cursor sizes and used a Fitts’ law task to analyze user performance. According to the results, subjects were slower and their throughput was lower when they hit a solid surface to interact with virtual targets. Our results show that Fitts’s model can be applied to these systems to measure and assess participant training.

CCS CONCEPTS

• Human-centered computing → Virtual reality; Pointing devices; HCI theory, concepts and models.

KEYWORDS

Virtual Reality, Fitts’ task, performance assessment, reaction test, mid-air interaction

1 INTRODUCTION

Since 3D immersive Virtual Reality (VR) systems have become more affordable and thus more accessible, many application fields have attempted to implement their tasks in VR, due to the controllability that a virtual environment (VE) affords. Compared to an analog system or a hardware setup, using a VE makes it easier for practitioners to collect data from the user. Furthermore, making changes to the VE is much more affordable and faster compared to real-world implementations.

One option to assess user performance is through performance assessment with psychometric tasks/tests, e.g., via the the Senaptec

2 PREVIOUS WORK


\[
\text{Movement Time} = a + b * \log_2 \left( \frac{A}{W} + 1 \right) = a + b * \text{ID} \tag{1}
\]

In equation 1, \(a\) and \(b\) are empirical constants, typically found by linear regression. \(A\) is the movement amplitude, i.e., the distance between targets, and \(W\) the target width. The logarithmic term represents the task difficulty and is called the index of difficulty, ID.

Effects of visual feedback: While 2D pointing and selection are well understood, 3D pointing and selection in VR are more involved and less studied. Research is still exploring methods and approaches that enhance user performance in VR. For instance, previous research on visual feedback showed that highlighting objects in VR increases the selection time and throughput while decreasing errors [25]. Also, environmental depth cues, such as shadows [17], motion parallax [24] and textures [12] help the user to perceive depth better, which increases selection performance. When we designed our experimental VE setup, we took all these results into account.

Passive Haptics: Passive haptics describes feedback which allows user to feel physical objects in a VE [19]. Such feedback increases the sense of presence in VR, since the user cannot see the real environment in VR. However, it improves user performance by providing haptic touch cues from the environment [15, 23, 27]. Previous research [6] showed that using a static surface while subjects are using an immersive VR headset improves user performance.

Cursor size variation: As mentioned in the introduction, VR provides a controllable environment and it is possible to collect user data with different target and cursor sizes in VR. These different sizes can have a significant effect on user performance [29]. Fitts’ law can analyze user performance with different target sizes [9].
Cursor size variation for selecting a target has been explored, e.g., through the “prince” technique [14] or a bubble cursor [10]. With the bubble cursor, the size of the cursor shrinks or grows to touch the target, which was also evaluated in VR [26]. In the VR implementation, the size of the cursor was constantly changing and created a distraction, i.e., while the cursor was smaller for closer targets, it was suddenly bigger for further targets.

3 MOTIVATION & HYPOTHESES

We investigate one of the tasks included in the “Sensory Station”, which consists of a vertical large touchscreen that displays a grid of targets. This station serves as a tool for human performance tests, and is typically used in athletic training. We then modified the software of the re-implementation to collect sufficient data to analyze the performance as a Fitts’ task.

Since previous work [6] showed that using a static surface while subjects are using an immersive VR headset improves user performance, we hypothesize that user performance significantly increases when a real-world object is used for passive haptics compared to mid-air interaction. Moreover, previous research showed that larger cursor sizes improve movement speed and both cursor and target size improve precision [29]. The issue of an optimal target and cursor size and how they affect user performance still remains. Our second, exploratory hypothesis is that cursor size improves user performance with passive haptic feedback.

4 USER STUDY

Participants: We recruited twelve participants (5 female) from the community. The average age was 26.33 ±3.7 years. All subjects were right-handed and used their dominant hand to execute the task.

Apparatus: We used a PC with i7-5890, 16 GB RAM, and RTX2080 graphics. We used an HTC Vive Pro headset and its controllers as input devices. To track bare hand interaction, we attached a Leap Motion to the front of the headset.

Procedure: After filling a demographic questionnaire, participants were informed that they would perform the experiment while standing. In the VE, subjects were placed in a virtual gym. They were facing a vertically-oriented plane where the targets were placed in a 6x6 grid of targets, with 8 cm spacing. The targets were push buttons, where the user has to ‘push’ a button to select a target, see Figure 1(e). To prevent participants from only “tenuously” touching targets, they had to be approached from their front and depressed by a pre-specified amount before they triggered, which makes the task more similar to other athletic performance assessment systems. The color of idle targets was set to gray. A black semi-transparent plane was placed behind the targets to improve contrast and ease the visual perception of the next target, but transparency was adjusted so users could still see through that plane, to get appropriate self-motion and depth cues from the VE.

Subjects performed the study under three different Interaction mode conditions. In the first, called VR Controller condition, subjects used a cursor placed 5 cm “above” the top of the VR controller. In this condition, subjects were physically facing into an open space, so that they could freely move their arms and hands.

In the other two conditions, users used their bare hands to interact with targets, through the Leap Motion attached to the front of the headset. For these conditions, we asked subjects to face towards a black wall with a surface that has a dense, thick pile of polypropylene on top (similar to a rug, but dampens sound). Normally, a Leap Motion cannot track hands coming into contact with a surface. Yet, our wall material allowed us to reliably detect hand poses even when the user placed their hand on the surface. We carefully calibrated the hand tracking so that physical motions matched virtual ones. We used the systems’ internal calibration methods and also verified distances and the positions in the real world. For this, we placed two 1 cm cubes at 40 cm distance into a VE. We then asked pilot users to place their index fingers into these boxes and measured the real-world distance to be 39.3 cm, which is very close to the virtual one. We used Leap Motion’s Orion SDK in see-though mode to verify that the virtual hands and hands’ images matched. In the second condition, subjects were asked to select the virtual targets placed on the physical wall by hitting them with their palm. In effect this condition implements a form of passive haptics [7, 13]. We call this the Haptic feedback condition. We carefully calibrated the target grid to match the wall. Four opaque calibration squares were placed at the corners of the grid and participants had to hit them. We then also rotated or re-positioned the target grid until the largest (virtual) cursor was not visible under the squares.

In the third condition, we asked subjects to perform the same task in mid-air, i.e., without touching the wall. We refer to this as the Mid-air condition. To ensure comparability, we placed a selection cursor 5 cm above the virtual hand in the second and third conditions. To aid with depth perception and help users to reliably see their (virtual) hand position, we used the hand models provided by Leap Motion, see Figure 1(d). In all three interaction modes, the height of the target button grid was adjusted, so that
each individual could easily perform the study without having to stretch.

Subjects were asked to select the current, yellow target. When the cursor touched a target, we highlighted the target in blue, to give users feedback that they are in contact with the button. When subjects “pushed” more, i.e., activated the button, we changed the color of the button to green to indicate a successful selection/“hit”. If the user missed the target or pushed the wrong one, we changed the color of the target to red to indicate a task error and played an error sound, but still recorded the location in space. Targets that were hit by accident did not change color. We also did not allow our system to re-select the same button as a target in a set of trials.

At the beginning of each set of individual movement trials, we randomly chose one of the 36 buttons as the first target. After selection of that button and to vary the ID in the experiment, the software then randomly selected the next one within a set of four designated Target Distances TD1: 16, 22.6, 24 and 32 cm. This continued until all the buttons were selected or there was no more available target left within the set of next target distances. This procedure also allowed us to define a clear end to the task repetitions. Whenever there were no more (valid) targets in the arrangement, we let users pause for 2 seconds to indicate the end of the repetition and then started the next set of trials. We also varied the cursor and target button sizes, with three different Sphere Sizes SSV1. When the target button width was 1 cm, we set the cursor width to 1.6, 3.2, or 4.8 cm or, vice versa, we set the target button width to 1.6, 3.2, or 4.8 cm when the cursor width was 1 cm.

If participants felt fatigued, they could take a break up to a minute between each repetition and up to 5 min. between interaction modes. The whole study took around 30 min. per participant.

4.1 Experimental Design
Participants experienced 6 main experimental conditions: three Interaction Modes (IM1: VR controller, Haptic feedback and Mid-air) and two Sphere Size Variations (SSV2: Cursor or Target), in a IM1 × SSV2 within-subject design. To vary the ID for Fitts’ law, we varied Sphere Sizes, see above. Conditions and Sphere Sizes were counterbalanced with Latin squares. Participants repeated each individual combination 3 times. Subject’s movement time (ms), error rate (%), and effective throughput (bit/s) were measured as dependent variables. Based on the different values for TD1 (which varied within the trial set) and SSV1, we evaluated 10 unique IDs between 1.94 and 4.39. Since the number of selected targets could not be the same in each repetition and trial, the number of collected data points varied slightly, but we collected on average between 30 and 31 data points in each repetition, for an average of ~540 data points for each interaction mode condition. In total, we collected 17254 data points for each movement time, error rate, and throughput dependent variables.

5 DATA ANALYSIS
The results were analyzed using (RM) repeated measures ANOVA with α = 0.05 in SPSS 24. We considered the data as normal when Skewness and Kurtosis values were within ±1 [11, 21]. We deleted “double click” data (%50.74 of the data), where the next target was selected without hitting another button. We used the Sidak method for post-hoc analyses. Below, we use *** for p < 0.001, ** p < 0.01, * p < 0.05, and n.s. for not significant.

5.1 Results
Time and error dependent variables were normal after log-transform and the throughput variable had a normal distribution. The one-way ANOVA results are shown in 1. For the interaction mode, Mauchly’s sphericity test was not violated for time (χ²(4) = 3.03, n.s.), error rate (χ²(2) = 2.144, n.s.), and throughput (χ²(2) = 2.592, n.s.). For the ID condition, Mauchly’s sphericity test was violated for time (χ²(4) = 100.42, p < 0.001) and throughput (χ²(4) = 129.72, p < 0.001), but not for error rate (χ²(4) = 48.18, n.s.). For the RM analysis, we used Huynn-Feldt correction, since $\epsilon = 0.334 < 0.75$ for time and $\epsilon = 0.283 < 0.75$ for throughput.

<table>
<thead>
<tr>
<th>Table 1: RM ANOVA results</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Movement Time</th>
<th>Selection Sphere</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>F(2,22)= 52.70</td>
<td>$\eta^2 = 0.748$</td>
<td>***</td>
</tr>
<tr>
<td>F(1,11)= 13.32</td>
<td>$\eta^2 = 0.548$</td>
<td>**</td>
</tr>
<tr>
<td>F(3.007,33.082)=102.57</td>
<td>$\eta^2 = 0.903$</td>
<td>***</td>
</tr>
</tbody>
</table>

| Error Rate | F(2,22)= 7.98 | $\eta^2 = 0.42$ | ** |
| Selection Sphere | F(1,11)= 5.00 | $\eta^2 = 0.313$ | ** |
| ID | F(9,99)= 10.73 | $\eta^2 = 0.494$ | *** |

| Effective Throughput | F(2,22)= 64.181 | $\eta^2 = 0.854$ | *** |
| Selection Sphere | F(1,11)= 0.18 | $\eta^2 = 0.016$ | n.s. |
| ID | F(2.53,28.00)=18.757 | $\eta^2 = 0.63$ | *** |

Figure 2: Time (a) error rate (b) and effective throughput (c) results for interaction modes.

Time: The results for time are shown in Table 1, and for interaction mode in Figure 2(a). Subjects were faster when they performed the experiment with a VR controller and also when we varied the target size (instead of cursor size).

Error rate: The error rate results are shown in Table 1 and for interaction mode in Figure 2(b). The error rate was higher when subjects had to interact with the wall and decreased when we used a fixed-size (1 cm) cursor and varied target size.

Throughput: The throughput results are shown in Table 1 and for interaction mode in Figure 2(c). Subjects throughput significantly decreased when they used the passive haptic feedback provided by the wall. Moreover, there was no significant difference between cursor and target size variation for throughput.

Interactions: We used two-way RM ANOVAs to detect interactions. For the interaction of interaction mode and sphere selection variation, Mauchly’s sphericity test was not violated for time ($\chi^2(2) = 0.13$, n.s.), error rate ($\chi^2(2) = 2.36$, n.s.), and throughput ($\chi^2(2) = 1.51$, n.s.). All dependent variables showed significant interaction between interaction mode and sphere selection variation conditions: time F(2,22)= 12.06 p < 0.001, $\eta^2 = 0.523$, error rate...
F(2,22) = 4.36, p < 0.05, \eta^2 = 0.248, and throughput F(2,22) = 7.92 p < 0.01, \eta^2 = 0.419. Subjects were faster with the small cursor (1 cm) and larger targets when passive haptic feedback was provided. Also, subjects’ error rate dropped significantly when they performed the task with larger targets with a small cursor (1 cm) in all conditions. Further, throughput increased in the mid-air condition when larger cursors were used with small targets (1 cm). On the other hand, subject throughput decreased in the haptic feedback condition when larger cursors were used with a small target (1 cm).

When looking at the throughput. (Figure 2) with (passive) haptic feedback. Yet, with a VR controller and haptic feedback conditions, users had to raise their hands and perform the experiment approximately at their shoulder level. Due to the different grip style of the VR controller, we were not able to raise the targets as high in this condition. This made it somewhat easier for users to execute the VR controller condition. The average task fatigue was 4.75 according to 7-point Likert scale results. This could be considered normal after participating for a performance assessment scenario. Moreover, we did not observe any motion or cybersickness that could affect users during the study.

A second potential explanation of our result is the visibility of the cursor and target buttons in VR. Previous research showed that controller size has a significant effect for pointing tasks [28], which makes it advisable to place the cursor above the VR controller. However, when the cursor is placed 5 cm above the user’s hand, the virtual hand can obscure some of the targets in the experimental setup. Curiously, in the interview at the end of the experiment, subjects said that they preferred to see virtual hands, as it helped them to more easily align their hand with the targets.

While these observations could potentially explain the performance increment for the VR controller, it still does not explain the difference between the mid-air and haptic feedback conditions. This is a very surprising result, as there is a commonly held assumption that a “perfect” haptics system would be part of an “optimal” VR solution. While vibration has not shown to be strongly beneficial [6, 22, 27], a haptics system that can stop the hand might perform better. Yet, with a task that involves hitting targets on a plane, our passive haptics wall “system” exhibited worse performance, even though it is sufficient to support the task.

We believe that participants had a hard time suspending their disbelief – they need to hit a wall, but can see neither their hand nor the wall, which may have slowed them down. Somewhat understandable, they were reluctant to hit the wall with full speed. The end-to-end latency of the system, on average 36 ms, was reasonable, but may have still played a role in our result. We believe that more research into the effectiveness of passive haptics is needed to identify the reasons behind our results.

When developers decide on the target and cursor sizes for a VR system like ours, we suggest to calibrate these sizes according to a performance evaluation with users. The results on various cursor and target size also supports previous research of Wang and MacKenzie [29], in that cursor and target size has a significant effect on user performance and should be considered as a variable during performance evaluation.

7 CONCLUSION AND FUTURE WORK

We identified that using a passive haptic surface for a psychometric task designed to assess user performance in VR might not be the best option; while conventional systems use this approach, subjects performance significantly decreased with a passive haptic surface in our VR system. Sensory stations are also used in rehabilitation and medical research [1]. In the future, we plan to collect data from the corresponding populations to further evaluate the potential of performance assessment through Fitts’ law in different applications.

ACKNOWLEDGMENTS

Thanks to NSERC for supporting this research project.
REFERENCES


