2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) | 979-8-3503-2838-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISMAR59233.2023.00122

AR Guidance Design for Line Tracing Speed Control

Jeroen Ceyssens* Expertise Centre for Digital Media, Hasselt University - Flanders Make, Belgium

Bram van Deurzen† Expertise Centre for Digital Media, Hasselt University - Flanders Make, Belgium

Gustavo Rovelo Ruiz‡ Expertise Centre for Digital Media, Hasselt University - Flanders Make, Belgium

Kris Luyten§ Expertise Centre for Digital Media, Hasselt University - Flanders Make, Belgium

Fabian Di Fiore¶ Expertise Centre for Digital Media, Hasselt University - Flanders Make, Belgium

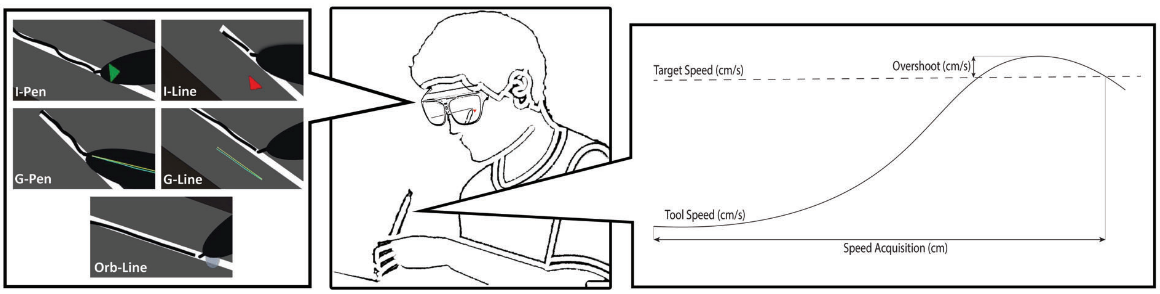


Figure 1: Visualizations tested within the study (left), the task users carry out (middle) and the metrics we primarily study (right).

Abstract

In many jobs, workers execute precise line tracing tasks; welding, spray painting, or chiseling, for example. Training and support for such tasks can be done using VR and AR. However, to enable workers to achieve the required precision in movement and timing, the effect of visual guidance on continuous movement needs to be explored. In VR environments, we want to ensure people are trained so that the obtained skill is transferable to a real-world context, whereas, in AR, we want to ensure an ongoing task can be completed successfully when adding visual guidance. To simulate these various contexts, we employ a VR environment to investigate the effectiveness of different visualizations for motion-based guidance in a line tracing task. We tested five different visualizations, including faster and slower arrows on the pen, the same arrows on the line, a dynamic graph on the pen or line, and a ghost object to follow. Each visualization was tested with the same set of five lines of different target speeds (2 cm/s to 10 cm/s in steps of 2 cm/s) with a training line of 5 cm/s. Our results show that the example ghost on the line turns out to be the most efficient visualization for allowing users to achieve a specific speed. Users also perceived this visualization as the most engaging and easy to use. These findings have significant implications for the development of AR-based guidance systems, specifically in the realm of speed control, across diverse domains such as industrial applications, training, and entertainment.

Index Terms: Human-centered computing—Visualization—Empirical studies in visualization; Human-centered computing—Visualization—Visualization design and evaluation methods

*e-mail: jeroen.ceyssens@uhasselt.be
†e-mail: bram.vandeurzen@uhasselt.be
‡e-mail: gustavo.roveloruiz@uhasselt.be
§e-mail: kris.luyten@uhasselt.be
¶e-mail: fabian.difiore@uhasselt.be

1 Introduction

Fine-grained motion-based tasks typically require users to carefully navigate an area with their hands while performing various manual actions. This is especially the case for line tracing activities (e.g., welding, caulking), where users must balance the tool accurately along a specific path at a certain speed without deviating from the target area. Depending on the activity, utilized tool, and surface area, the tool speeds required in these activities can vary. The tool speed is typically also based on a metric called the "feed rate" (the rate at which material "feeds" out of the tool). Speed control and accuracy are vital in balancing the feed rate and the tool speed to achieve proper end results. For example, in welding, the torch speed of a professional welder can typically range from 2 cm/s to 6 cm/s, which also depends on the feed-rate set on the welding tool [7].

Training and support for motion-based tasks can be done using Augmented Reality (AR) and Virtual Reality (VR) [5, 6, 23, 24]. However, to enable workers to achieve the required precision in movement and timing, the effect of visual guidance on continuous movement needs to be carefully explored. So AR guidance visualizations must be precisely crafted to avoid a negative impact on task execution. The additional information provided during the execution of a motion-based task may briefly interrupt or distract the user, as that information needs to be absorbed and interpreted, which in turn can affect the accuracy and speed of the fine-grained movements, ultimately leading to a decline of the quality of the end result.

Our goal is to provide insights into which types of visualizations are most effective for guiding users through fine-grained line tracing tasks that require precise motor movements and speed control. By doing so, our objective is to enhance the design and implementation of AR guidance methods for various motor skill activities. Figure 1 showcases the measurements that are of primary interest to our study. This includes the variation in users' speed from the target speed across different visualizations, the distance traveled before they achieve the correct target speed, and the extent to which users exceed the required speed when trying to achieve it.

To properly analyze these concepts, we conducted a user study that explores five different visualizations for line tracing activities on top of a surface. The study is conducted within VR, where we

DOI 10.1109/ISMAR59233.2023.00122

analyze the impact of instructions during a line tracing activity with a pen (see left of Figure 1). Given that virtual reality enables us to simulate occlusions commonly encountered in augmented reality, we anticipate that the insights gained from our findings can be applied to AR guidance scenarios.

This paper presents the following contributions:

- We explore the impact of visual guidance on the user performing line tracing tasks that require high precision.
- We compare three “low information density” visualizations for providing guidance on a tracing task.
- We identify the speed differences between the target speed and the actual conducted speed caused by the provided guidance, thus mapping the possible error that may occur.

To this end, we performed a comparative interaction study. Our results show that using demonstration for guiding the user makes them achieve the correct speed sooner, allows them to trace the speed more accurately, and gives an overall better usability experience. However, this is at the cost of a sharp increase in speed at the start of the task. The implications of our findings are substantial for the development of augmented reality guidance systems, particularly in the domain of speed control, with wide-ranging applications in industries, training programs, and entertainment.

2 Related Work

This paper takes knowledge from three main categories of research, human processing during motion, XR for motor skill tasks, and XR guidance during motion. We present past works and literature from within these categories.

2.1 Human-Processing during Motion

When aiming for a particular speed during motion, there may exist a potential trade-off between the speed and the accuracy of the movement [14]. We intend to balance this trade-off by guiding the users toward the correct speed while allowing them to keep the accuracy of their movement. Nonetheless, Maeder et al. [12] discussed that humans have a limited capacity for processing visual information (our "cognitive bandwidth") and that interface design can significantly impact user understanding and performance. Their model imposes an existing limit on the understanding of visual interfaces and emphasizes that using the power of abstraction allows us to present data in a way that assists faster human understanding. Thorpe et al. [16] also previously found that humans have a limited reaction time to process visual information when appearing at high speed. Similarly, Liu et al. [11] found that high latency in interactive visualizations reduces the rate at which users make observations and leads to lower performance even at subsequent lower latency visualizations. These studies emphasize the importance of interface design in facilitating effective visualization comprehension and user performance. Low latency and abstraction are important factors to consider in designing effective visualization interfaces for human performance.

2.2 XR for motor skill tasks

XR technology has repeatedly been shown to be an effective learning method compared to video-assisted instructions [1, 4] or traditional learning methods [19] for motor skill learning. When looking at motor control skills, it has also been shown to improve performance and learning of athletic activities [9] and in rehabilitation [25]. For guidance during activities, Yang et al. [21] studied the effects of employing the "ghost" metaphor in VR for guidance in calligraphy, where it was found that the VR guidance achieved equal performance to traditional methods for less training effort. Nomoto et al. also tested how visuo-haptic interaction with a dummy hand as a "ghost" affects the accuracy and speed performed while drawing and found a positive improvement in the accuracy but a decrease in the speed performance [13]. Meanwhile, Yi et al. studied how visualizing trails for object interaction with controllers and hands can affect the fine motor control ability within VR [22]. Wierinck et al. [20] studied augmented feedback for dexterity skills, specifically with dental students. They found that those receiving augmented feedback outperformed those without it in the performance test. Buchner et al. [3] also studied AR’s effects on the cognitive load and performance of motor and cognitive tasks. They found that AR has the potential to decrease cognitive load and increase task performance, primarily in-situ spatial AR. However, they mentioned that see-through AR could gain the same benefits with the proper value-added design of instructions. In general, these works imply that implementing AR and VR technologies can enhance training and learning experiences in various domains and that careful attention to cognitive load and interface design can help optimize user performance and understanding.

2.3 XR guidance during motion

Within our paper, we specifically focus on guiding the user toward controlling their speed of motion for specific line tracing activities. There has been prior work regarding guidance in AR or VR for following specific paths. Peternel et al. [15] used projected visualizations on top of the hand to guide the user’s hands toward a specific target area. They tested 2D and 3D visualizations and found that the AR-guided movement increased user accuracy by around 85%, compared to simple video instructions. In a similar study regarding arm movements, Volmer et al. [18] used precues (visualizations that show beforehand what the next steps entail) to guide the user for specific procedural tasks. For their example cases, showing a simple line towards the next target area proved to be the most efficient visualization. Similarly, Liu et al. [10] tested within VR how different styles of precues affect the efficiency and accuracy of users performing directed motions along a specific path with their arms. Yu et al. [23] further explored how such cues can direct users during full-body motions. They suggested several motion visualizations to direct the user using first-person (user perspective) and third-person (using an avatar) methods. They found first-person instructions to be the most efficient. One common visualization we often encounter within these works is using example movements to guide the user. Little information, however, is present on how such example visualizations help users in achieving the correct speed during motion since most works focus on performing the motion as fast or accurately as possible.

3 Design of Speed Guidance for Line tracing

The primary focus of our research is to design AR guidance visualizations for situations where users need to trace a line on a surface area at a specific speed. Some real-life examples of these situations include gluing, welding, painting, and caulking. We define three main aspects we employ in our design: visualization location, user adaptation, and continuity. These are identified based on prior works within the fields of XR guidance for motion and motor skill tasks. Our goal is to test the effects of these aspects on guiding users towards a correct speed, which is primarily still unexplored.

3.1 Visualization Location

Since point activity focus is important within these activities, visualizations should try not to obscure the activity, while still remaining close enough to ensure proper interpretation. Within our use case, we consider two locations that can be utilized for these visualizations: (i) on the line where the activity takes place (similar to [18]), and (ii) the tool used to carry out the activity (like in [1]).

3.2 Discrete Instructions

Line tracing activities often require high-precision motion, which makes it essential for the visualizations to incur the least amount of cognitive load. Since the cognitive load is highly dependent on the amount of information that needs to be processed at a given time, we consider the continuity of the information (when and what information is given). Within our use case, we look at instructions that are discrete or continuous in nature. Discrete instructions are clear, individual steps that need to be taken to reach a desired outcome. For example, in an AR/VR application, discrete instructions can be provided as textual or visual prompts, such as pop-up windows, tooltips, or arrows, which indicate a singular target activity [10, 18]. Without user adaptation, an example of line tracing motion would be the text prompt "Move along this area at a speed of X cm/s". This instruction is hard to understand for novice users and is less interesting within our context. To include user adaptation within this type of instruction, we consider using a discrete arrow-based glyph which indicates the need to either increase or decrease the user's current speed. Mapping the context of increasing speed to a positive relation (+) and decreasing speed to a negative relation (-), we can shape the glyph of a desired speed increase to a green upward arrow (▲) and desired speed decrease to a red downward arrow (▼). Based on the location, this instruction presents us with two visualizations to consider for the user: (i) arrow glyphs on the tool and (ii) below the point of interaction.

3.3 Continuous Instructions

Unlike discrete instructions, continuous instructions involve ongoing guidance for maintaining a desired outcome and typically require continual monitoring and adjustments.

3.3.1 User-Based

A real-life example of continuous, user-based instruction is the speedometer inside a car, which provides a range of speeds and indicates the driver's current speed. This same concept was also used by Ipsita et al. [1] for providing guidance during welding training. As a speedometer includes more detailed information that needs to be processed (e.g., textual information, circular outlines, tick marks), we aim to simplify the design of our continuous, user-based instruction by turning to the concept of *Sparklines*: simple, word-sized graphics introduced by Edward Tufte [17]. These *Sparklines* often consist of heavily simplified graphs without any labels or tickmarks since their primary focus is the comprehensibility of data trends. We base ourselves on this design as it supports a visualization that can be interpreted fast, does not obscure the task at hand, and represents the user's current state compared to the desired state. For every single point in time, users can see the representation of their speed and the target speed to reach ( ). This visualization can also be presented on both the pen and below the line.

3.3.2 Task-Based

Finally, we have incorporated the capability to visualize the desired speed independently of the user's current speed. We do this by utilizing the concept of a demonstration object or "ghost" [13, 21], which represents the tool positioned atop the area and moves at the correct speed. Within our context, we generalize this "ghost object" to an orb on the area that moves at the desired speed ( ), that users can choose to follow directly or copy from a distance at the correct speed. By limiting the size of the visualization, we avoid occluding too much area with unnecessary details. However, due to the nature of continuous instructions that are not user-based, this visualization method is hard to represent the tool of the user without restricting the orientation and depth of the tool and thus occluding a large area. Hence we only focus on visualizing the "ghost" object on top of the area.

3.4 Design Overview

Through this design study, we have identified three viable visualizations that can be considered for guiding users toward the correct speed (seen in Table 1). Our goal is to compare these visualizations with both quantitative and qualitative data to understand their effectiveness better.

Table 1: The visualization designs mapped to continuity, location, and user adaptivity.

Visualization	Continuity	Location	User-based
Binary indicators (▼▲)	Discrete	Area/Pen	Yes
Speed graph ( )	Continuous	Area/Pen	Yes
Ghost demonstration ( )	Continuous	Area	No

4 User Study

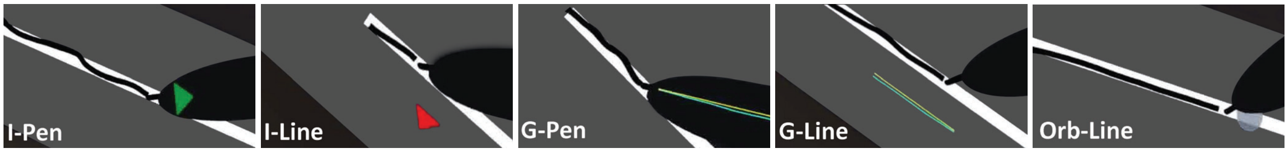
Through the systematic user study described in this section, we aim to uncover the effects of the previously described visualization techniques, within the context of a line tracing task, on user performance and perception.

4.1 Experimental design & Implementation

Our design study resulted in three visualizations with two variations in the first two cases (see Section 3), creating a total of five designs to explore as independent variables (IV) of the study. We name these I-Pen (Figure 2a), I-Line (Figure 2b), G-Pen (Figure 2c), G-Line (Figure 2d) and Orb-Line (Figure 2e). We test these visualizations by tracing a line with a pen. For the location, we place the instructions in the middle of the pen above the nozzle (Figure 2a and 2c) and below the intersection point of the pen with the line (Figure 2b and 2d). Visualizations always face the users, allowing them to focus on speed and position accuracy without rotating the pen and ensuring that the instructions do not occlude the drawing area. For the indicators, we enforce a threshold to recognize the correct speed: the visualization is hidden when users reach the target speed within a 20% threshold. This is to mitigate the impact caused by human-error margins. A 20% tolerance was selected to ensure fair representation across target speeds, as a static tolerance value (e.g., 0.1 m/s) compromises the balanced study design (apparent from pilot studies). The graph-visualisation we used in the study (Figure 2c and 2d)) uses a time frame of 1 second, allowing for interpretation and reaction time. The visualization's height is limited to half the thickness of the pen (1 cm) to avoid occlusion of the line. The orb visualization (Figure 2e) starts at one end of the line. When users enter the orb with the pen, it moves at the desired speed toward the other end of the line. Users are requested to match their pen speed to the orb's, and once the orb reaches the end of the line, it disappears automatically.

4.2 Procedure

In most example situations of real-life line tracing (e.g., welding, gluing, caulking), there is a feed rate present that controls the amount of material flowing onto the surface. Since the amount of feed rate in a material can directly affect the required speed of users and may influence the results significantly, we choose to conduct our study in the context of a simple drawing task in VR where the feed rate is constant. We utilize VR to simulate the drawing task and study the effects of the visualizations since it allows us to control all information perceived by the user and remove potentially unwanted variability. Within this study, users will hold a 3D-printed pen (Figure 3b) to trace a specific target line of 60 centimeters. The goal is to follow a predetermined line at the correct speed with this pen. To draw conclusions about the visualizations for low and high target speeds, we test the visualizations for a range of target speeds starting from 2 cm/s to 10 cm/s in steps of 2 cm/s. We omit variability within



(a) (b) (c) (d) (e)

Figure 2: The visualizations mapped to the study use case (background adjusted to a lighter shade for visibility) with (a) the indicators on the pen, (b) indicators below the line, (c) graph on the pen, (d) graph below the line, and (e) demonstration orb on the line

our study regarding tool familiarity since we assume most people within our target audience have gained prior experience holding a pen.

4.3 Participants

Within our test setup, the order of the visualizations is randomized across participants based on a 10×5 Balanced Latin Square (twice in row length to balance odd numbered visualization count), to minimize learning effects. A total of 20 participants (12 male, 8 female) with ages 20 to 52 (M=32.55, SD=9.73) were recruited to perform this study. Seventeen of the participants were right-handed, and three were left-handed. Participants were allowed to use the visualizations in their most dominant hand for this study. Users were also prompted to apply the "think-aloud" protocol and mention anything about the visualizations that would stand out to them.

4.4 Measures

To comprehensively evaluate the impact of various visualization techniques on user performance and perception during the line tracing task, we employed a range of measures to capture key aspects of participant behavior and performance. Some of the speed-related measures can be seen on the right of Figure 1. To make the comparisons, we accumulated all the readings to get one mean result for every line that was drawn. Hence for every quantitative measure, there is one result per participant (20), target speed (5), and visualization (5), with a total sample size of 500 (20x5x5). Within our study, the following dependent variables (DV) were analyzed:

Drawing speed One central measure of interest for performance was the speed at which participants traced the target lines. To decide the performance of drawing speed, we calculate the difference between the speed of the pen and the target speed at the time for every visualization. Here we establish our first hypothesis H_1*: The drawing speed differs across the visualizations.*

Achieving target speed For another performance metric, we examine the point along the line at which participants successfully first achieved the target speed. We consider a particular target speed as "achieved" when the user balances the speed first above the target speed and then lowers it back. Note that for this measurement, we only take into account the initial instance when the speed is acquired. For this DV, we establish the hypothesis H_2*: The position at which users achieved the correct target speed differs across the visualizations and target speeds.*

Speed overshooting The degree to which participants exceeded the target speed before achieving it correctly, as error rate. This variable offers insights into how participants controlled their speed based on the different visualization cues. Since line tracing activities can require high-precision speeds, overshooting can cause problematic situations where users are temporarily faster than what is allowed for that specific activity. To calculate the amount of overshooting caused by the visualizations, we measure the highest value of the tool speed from the start of the line until the target speed is first acquired. When the target speed was not acquired, the speed never reached the correct value, which implies no overshoot was present. We formulate the hypothesis H_3*: There is a difference in speed overshooting between target speeds and visualizations.*

Drawing & tool displacement We are interested in seeing the total displacement of the tool compared to the line as an error rate since users were required to balance the tool while acquiring the correct speed. For every single drawing position, we determine the distance to the projected point in the middle of the line. We are also mostly interested in seeing the total distance away from the center, which means we take the absolute value of this distance (negative values will not balance out the positive). Note that this position does not consider the depth of the tool since the drawing position is always aligned based on the depth of the line. We will look at the difference in depth of the tool compared to the projected point on the line to be drawn as a separate metric. We formulate hypothesis H_4*: There is a difference in drawing and tool displacement between visualizations.*

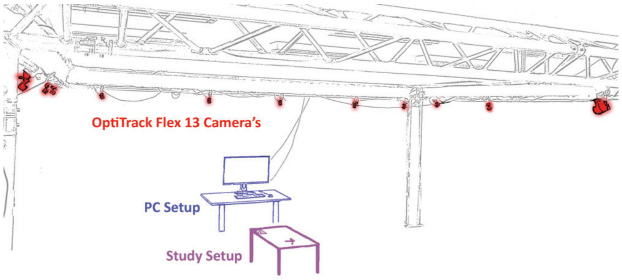
Perceived usability We collect several measurements to gauge participants' perceptions of usability. Users have to rate every visualization with the System Usability Scale (SUS) and NASA-TLX questionnaires after completing all lines for a particular visualization. We also questioned for every line they have to draw how they perceived the drawing accuracy, speed accuracy, distraction caused by the visualization, and overall understandability of the visualization. Finally, at the end of the study, we asked the participants for a final preference ranking of the visualizations. For this measurement, we formulate the final hypothesis H_5*: There is a difference in perceived usability between visualizations.*

4.5 Apparatus

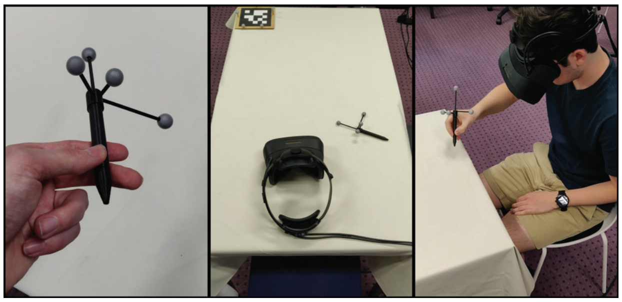
Within our study, we make use of the Varjo XR-3[1] headset. This XR headset is a VR headset that contains built-in AR passthrough capabilities and has a 70 ppd (pixels-per-degree) resolution. The Varjo XR-3 uses SteamVR lighthouses for outside-in tracking, however, since we require high-accuracy tracking for the pen, we make use of 12 OptiTrack Flex 13 camera's[2], which provide sub-mm accuracy of the 3D position at 100 FPS (layout of the cameras can be seen on Figure 3a). To calibrate the OptiTrack environment with the Varjo system, we utilize a Varjo marker (an inside-out based tracking method from Varjo) shown on Figure 3b with several OptiTrack rigidbodies attached to. The rigidbodies on top of the Varjo marker frame have been calculated and measured on the sub-mm level to align the two environments together. At the start of the study, the Varjo marker is scanned several times, and the position and rotation are aligned with the OptiTrack environment. This also allows us to

[1] Varjo XR-3:https://varjo.com/products/xr-3/ Last Accessed: 16/08/2023.

[2] OptiTrack Flex 13: https://optitrack.com/cameras/flex-13/ Last Accessed: 16/08/2023.



(a)



(b)

Figure 3: Practical setup of our study environment with (a) the OptiTrack environment and (b) the tools of the user study, including a 3D-printed pen, a Varjo XR-3 with Varjo Marker, and an example user carrying out the study.

use a lightweight 3D-printed pen as a tracking tool (seen on the left of Figure 3b).

4.6 Task

The digital platform and lines that appear within the study are positioned 10 cm higher than the physical table in front of the user to force users to keep the pen floating instead of directly touching the table. Touching the table would cause additional friction depending on the pressure put by the pen, which makes achieving higher speeds more difficult (this became apparent during pilot studies). To properly test the designs for line tracing activity, each participant had to carry out the following tasks:

1. First, participants must fill in informed consent, listing the details recorded within the study and approving the data usage. Afterward, every participant starts with an introduction phase of two lines they can trace at their own pace.

2. They then receive their first visualization and an introduction line to test the visualization (the introduction line is set to 5 cm/s). When the line is finished, it will turn green to confirm they finished it, and users are then given four questions they can answer, from 1 to 5, by using the pen (see Figure 4). These questions include "How would you rate your drawing accuracy?", "How would you rate your speed accuracy?", "How distracting was the visualization?", "How understandable was the visualization?". For all questions, except for question three, a higher value indicates a positive correlation.

3. After the training line, users are required to draw five lines at varying speeds (2 cm/s to 10 cm/s in steps of 2 cm/s), followed by the same four questions after every line (Figure 4). Whenever they are faced with a line, they are asked to stay as close as possible to the middle of the line while keeping the speed desired by the system. Users can also trace the line from left to right only, to prevent inadequate comparison of the visualizations.

4. After all five lines have been completed, the user will take a break from using the headset and will be asked to fill in a questionnaire containing the SUS questionnaire [2], and the NASA-TLX questionnaire [8], to rate the visualization they just experienced. We specifically tell them to rate the visualizations in the context of AR instructions in real-life use cases with example situations given (e.g., applying silicone at the correct speed).

5. Steps 3 and 4 are then repeated for each individual visualization (a total of five times for all conditions). The order in which the speeds appear within the lines is decided based on a 5x5 Latin Square.

6. After the final visualization has been explored, and the questionnaires are filled in, users are then prompted to rank the five visualizations from 1 (best) to 5 (worst), and also give a motivation for this specific ranking.

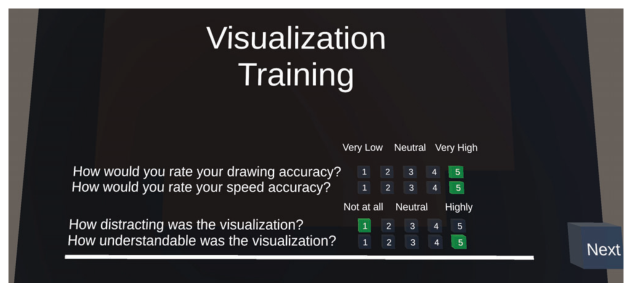


Figure 4: Questions asked after tracing each line; a rating can be given from 1 to 5.

5 Results

Our study primarily focuses on the metrics shown in Figure 1. That is, the general speed difference between the target speed and the actual conducted speed, the distance at which the target speed is achieved, and the amount of overshooting upon reaching the correct target speed. Aside from the speed metrics, we are also interested in the users' perceived usability of the instructions and the magnitude of displacement the instructions might have caused.

To equalize the results across all participants (i.e., the same amount of data points across the line), we split the target line of 60 centimeters into segments of one millimeter and calculate the mean tool speed of each participant within these segments (this ensures 600 data points for every participant, one for each millimeter segment), before calculating individual measures.

5.1 Speed Difference

A representation of the mean speed for every visualization (per centimeter for clarity) can be seen in Figure 5. At first glance, we notice that as the target speed increases, the difference in speed between the visualization methods also increases (particularly between the Orb-Line and all others and between the indicators and the graph). Meanwhile, for target speed 2 cm/s, the G-Pen and G-Line visualization speeds tend to stay above the target speed, while for target speeds above 4 cm/s, they tend to stay below. When running the Shapiro-Wilkinson normality test for all speed differences, we find that the I-Pen, I-Line, and Orb-Line are not normalized under all target speed conditions. Since not all visualizations contain normalized data, we test for significance using the Friedman rank sum test

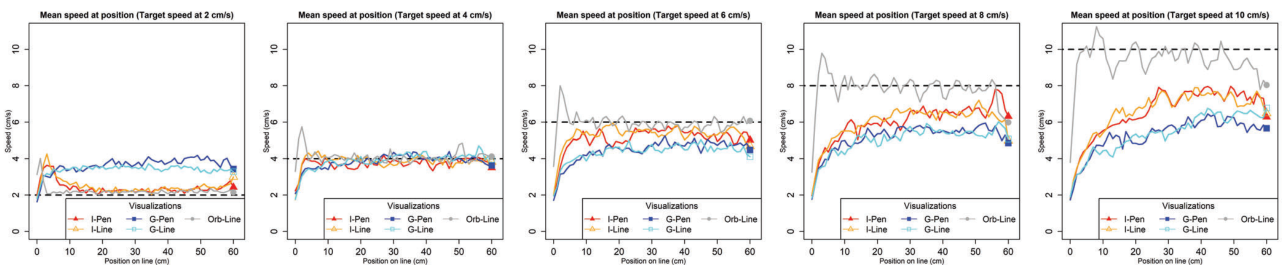
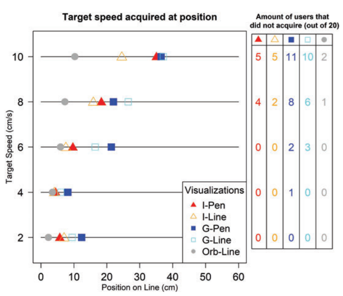
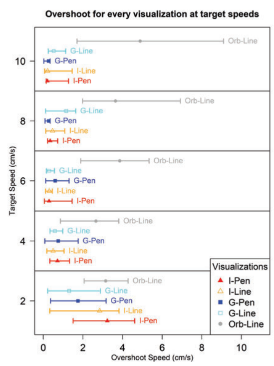


Figure 5: Plot of the average speeds obtained at each centimeter along the line for every visualization, categorized by different target speed conditions (represented by the dashed line).



(a)



(b)

Figure 6: Plots for all target speeds and visualizations of (a) the location on the line the target speed had been acquired (right side is the amount of users that did not achieve the target speed), and (b) the amount of overshooting when acquiring the target speed.

and use the pair-wise Wilcoxon post-hoc test (with Holm correction) to get the effect sizes. The complete results of the tests, including the effect sizes, are written in Table 2 in Appendix A. In general, the Friedman test reveals a significance in speed difference in all the target speeds. Comparing the visualizations individually, we find that target speed 4 cm/s is the only target speed with no statistical difference between the visualizations. We also see an overall significance between the Orb-Line and all other visualization methods. Consequently, we reject the null hypothesis of H_1 for the combination of all target speeds and for all target speeds not equal to 4 cm/s, which implies the results are statistically significant. For the 4 cm/s, we accept the null hypothesis of H_1, indicating no statistical difference in speed between the visualizations at target speed 4 cm/s.

5.2 Achieving Target Speed

Figure 6a shows a mapping of the distance where the target speed is achieved. Note that for each target speed, there have been times when users were not able to reach the speed correctly (the right column of Figure 6a shows the number of participants). The data entries in which the user did not achieve the correct speed for a specific visualization were excluded from the plot. This was done to avoid the inclusion of outliers that could arise depending on the distance at which these entries would be placed. We can notice that in the case of 2 and 4 cm/s target speeds, the visualizations had very similar performances. However, at all higher target speeds, the gap in distance between the visualizations continues to increase. In particular, the Orb-Line visualization is continuously first for every target speed condition, which corresponds to the results seen in Figure 5.

When testing the acquired position data of each visualization for normality, for the target speed of 10 cm/s, all visualization conditions are normally distributed. For 2, 4, 6, and cm/s, at least one of the visualizations is not normally distributed. For all target speeds other than 10 cm/s, we conduct Friedman's test to check for significance within the data (since not all visualization methods are normally distributed). This reveals a statistical significance on the acquired position data for the target speed of 8 cm/s (Table 3 in Appendix A shows the complete results with the post-hoc Wilcoxon tests), and barely not for 6 cm/s (χ^2=9.05, p=0.05986). The post-hoc tests for target speed 8cm/s show a significance between the Orb-Line and all other visualizations. All calculated effect sizes for these significant statistics were large.

Since we can assume normality for all visualization conditions on 10 cm/s target speed, we also test the sphericity (using Mauchly), which showed that the data was spherical (p=0.88595). We run the one-way repeated measures ANOVA for target speed 10 cm/s and choose to set the missing values (participants who did not achieve the speed) to the end of the line (distance of 60 cm), where the results imply a highly significant difference. To get a full overview of the differences within the 10 cm/s target speed acquisition, we run the pairwise t-test. The full results of the ANOVA and t-test can be seen in Table 4 in Appendix A. This yields a significant difference between the Orb-Line condition and all other visualizations and a significant difference between the I-Line condition and the G-Line or G-Pen conditions. We reject the null hypothesis of H_2 for target speeds 8 cm/s and 10 cm/s, which implies a significant difference in when the target speed is acquired. However, we accept the null hypotheses of H_2 for the target speeds 2, 4, and 6 cm/s, meaning no difference in acquiring the target speed.

5.3 Overshooting

The median and the first and third quartiles for all the overshooting measurements per visualization and target speed can be seen in Figure 6b. At first glance, we notice the Orb-Line condition to have a higher overshoot for conditions 4, 6, 8, and 10 cm/s.

When conducting the Shapiro-Wilkinson test for the overshooting, we could not identify any target speed for which all visualization entries are considered normally distributed. As a result, we conduct Friedman's test to check for significance within the overshooting data for every target speed. All target speeds had a significant difference in overshooting between visualizations, so we ran the Wilcoxon test for every target speed and for all target speeds combined (full results found in Table 5 in Appendix A). For target speed 2 cm/s, there are no significant differences found between the individual visualizations. For 4 cm/s there is a significant difference between the Orb-Line and the G-Line conditions. For 6 cm/s, 8 cm/s, and for the overall comparison, there is a significant difference in overshooting between the Orb-Line condition and all other visualizations. For 10 cm/s, there is a significant difference between the Orb-Line condition and the I-Line and I-Pen conditions.

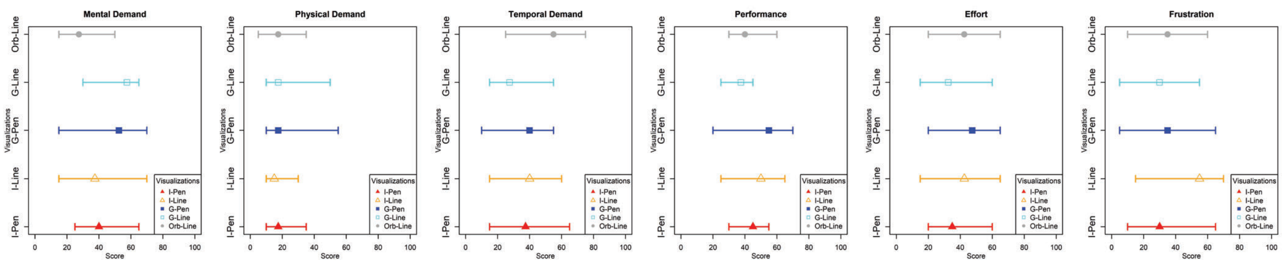
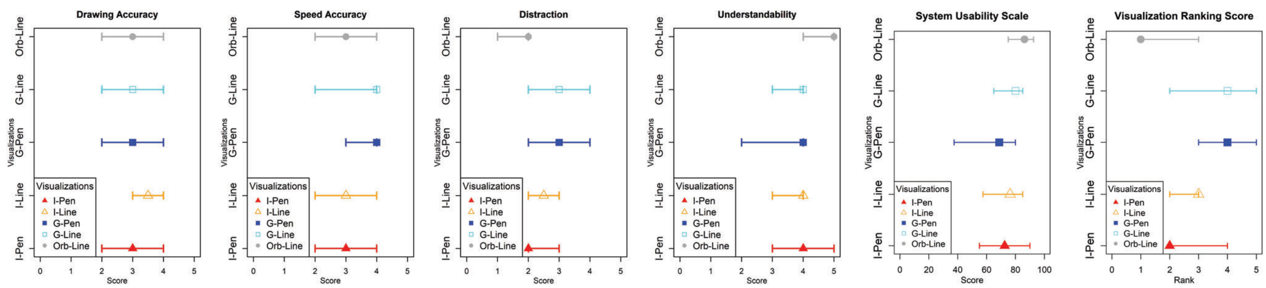


Figure 7: Median, first quartile and third quartile of the NASA-TLX results from the participants across all visualizations.



(a) (b) (c)

Figure 8: Median, first quartile and third quartile showing for each visualization (a) the results of the Likert-scale questions asked in the study, (b) the scores of the System Usability Scale (on 100), and (c) the ranking score of each visualization (1=best, 5=worst).

All the effect sizes found for the overshooting differences can also be categorized as large effect sizes. As a result, we can safely accept the null hypothesis H_3 for target speed 2 cm/s, and accept the alternative hypothesis H_3 for target speeds 4, 6, 8, and 10 cm/s, but also for the combined target speeds.

5.4 Drawing & Tool Displacement

After running the Kolmogorov-Smirnov test for the drawing and tool displacements, we find that all data can be assumed to be normalized for the drawing displacements but not for the depth of the tool. We use repeated-measures ANOVA to calculate the statistical significance of the drawing displacements. For all the target speeds, we find no statistical significance, which implies the drawing displacement to be non-significant across the visualizations (lowest p-value was 0.0537 for target speed 8 cm/s). Also the overall drawing displacement across all target speeds was non-significant (p=0.0611). For the tool depth, we carry out Friedman's rank sum test and found no statistical significance between the visualizations for any target speed (all p-values > 0.14). Since there was no statistical evidence for either the drawing distance or the depth of the tool, we can safely accept the null hypothesis of H_4 that there is no difference in drawing and tool displacement between the visualizations.

5.5 Perceived Usability

The results of the NASA-TLX can be seen in Figure 7, and the results of the SUS questionnaire are in Figure 8b. For the SUS questionnaire, the Orb-Line had the highest perceived usability score (M=86.25, sd=15.57). After the study, we also asked the participants to create a ranking of the visualizations from best to worst (1=best, 5=worst). The ranking results can be seen in Figure 8c, where the Orb-Line condition had the best rating (M=1, sd=1.56) and was selected first twelve times. Interestingly, the G-Line condition was selected first four times (second place in terms of the amount of first picks), with user testimonies explaining that it was the most calming out of all the visualizations. We can see this relation expressed in the SUS as well, where the G-Line has the second-highest score (M=80, sd=16.91). However, this visualization was also picked last the most (together with the G-Pen condition, both seven times).

Running the Kolmogorov-Smirnov normality test on the SUS scoring, we notice the dataset to be considered normally distributed. We test the SUS score dataset on sphericity using Mauchly's Test for Sphericity, which implied that sphericity can not be assumed (p=0.04604). To adjust for the sphericity, we run the Greenhouse-Geisser and Huynh-Feldt Corrections (ε=0.67248), which corrects our SUS scores and implies a significant difference (F(4,76) to F(2.69,51.11), p=0.0038). For the NASA-TLX results, we found that the distribution of questions was not normally distributed. After testing the significance of the question results with Friedman, we find a significant difference in the "Mental Demand" and the "Temporal Demand" questions. However, after running the Wilcoxon test, we find no significance between the individual visualizations for the two questions (full results see Table 6 in Appendix A).

The rankings and the scores for the different questions also appear to be not normally distributed (according to the Shapiro-Wilkinson test). As a result, we use the Friedman rank sum test to test for significance. We find no significance for all the questions except for question 4, "How understandable was the visualization?". We also find a statistical significance for the ranking between the visualizations using Friedman. After running the Wilcoxon test for both measurements, we find a significant difference for question 4 between the Orb-Line and the I-Line, G-Pen, and G-Line conditions and no statistical difference between the visualizations for the ranking (full results see Table 6 in Appendix A).

Based on the statistical significance we found in the overall NASA-TLX, overall ranking, SUS questionnaire (individually), and question 4 regarding "understandability", we can safely reject the null hypothesis of H_5, and accept the alternative hypothesis H_5, implying a difference in perceived usability between the visualizations.

5.6 Participant Testimonies

Since users were asked to apply the "think-aloud" protocol, several comments were made throughout the study. For the indicators (I-Pen/I-Line), participants mentioned the switching between red and green (caused when being unstable around a target speed) to be annoying during the activity. Half the participants also mentioned that they would only use the colors of the indicators to understand what should happen, not the orientation of the arrow. Participants also mentioned the graph visualizations (G-Pen/G-Line) to be distracting since they had to look out for small adjustments within the visualization. The Orb-Line had the highest ranking, with participants mentioning it was "very easy and intuitive, you just have to follow it". However, users that did not prefer the visualization mentioned that "the start goes too fast to follow, so it becomes hard to keep up", which is a problem specifically occurring at higher target speeds.

When the visualizations requested higher target speeds, participants would mention that they physically could not achieve this speed but would still achieve them when the Orb-Line visualization showed it. Almost all participants (except for two) would also try to catch up to the orb the moment it leaves, causing an initial increase in speed and allowing them to follow the speed afterward. According to many participants, the other visualizations allowed them to be more relaxed during the activity since they would always consider the user's current state instead of forcing them into one, which gives them more flexibility. The graph visualizations (G-Pen/G-Line) were specially mentioned to be forgiving since users felt at times that it was okay to have the lines close together but not perfectly matching, causing them to stay at a more comfortable pace. Interestingly, question two on Figure 8a also shows that participants felt they achieved the correct target speed, which could be related to the visualization giving lower pressure to perform.

6 Discussion & Future Work

The goal of our study was to explore the effectiveness of different visualizations for motion-based guidance in a line tracing task. Specifically, our study aimed to investigate the effectiveness of five different visualizations for speed control in line tracing and to identify the most efficient visualization for this task. As seen in section 5, there have been several significant findings during our study. The Orb-Line condition has repeatedly been shown to be the most efficient at making users acquire the correct speed, keeping the correct speed, and in terms of overall usability. The findings of our study gave an indication that when the visualization does not adapt to the user performance for guiding them toward the correct speed, better results could be acquired. From the testimonies of the users, they mentioned that the other instructions gave them "less pressure to perform", while the Orb-Line instruction was more demanding to the users. The results could also be in part due to the Orb-Line visualization giving "more clarity" to what is expected. Future studies will have to confirm these findings and see how much user performance adaptation is required.

Although the Orb-Line visualization outperformed all other visualizations, there are still a few problems with this concept, mainly due to the overshooting it causes. Since user reaction is always delayed, at a fast target speed, the orb will have traveled a further distance. This causes many participants to go way past the target speed at the start of the line and incurs a significant speed overshoot (seen in the significant results). To avoid these issues, there are methods that could be considered for future work. One method would be to let the orb start together with the user and gradually move it toward the target speed based on the general practices of the activity. Another option is to introduce a countdown for the start of the activity and preview the orb speed once before the actual execution, to inform users beforehand of the correct target speed. These suggestions should still be confirmed and explored further in future works.

For the NASA-TLX results and the ranking, while a statistical difference was present across all visualizations, no significance was found between individual visualizations. This implies that the exact relations need to be explored further in future studies, potentially with a higher sample size, to verify whether there are differences to be found. We also noticed that the indicators (I-Pen/I-Line) outperformed the graph visualizations (G-Pen/G-Line) in terms of speed difference for 2 cm/s, 8 cm/s (only I-Line), and 10 cm/s. One could assume that, for the design of the instructions, showcasing continuous data might be less valuable compared to directly indicating the action required, because users found the graphs too distracting. However, other visualizations would need to be explored to confirm these findings (Orb-Line was also continuous and more preferred). Making the binary indicators (I-Pen/I-Line) more gradual (e.g., increasing or decreasing the arrow sizes based on how much adaptation is required) could give different results for the set continuity. While slightly more users preferred the conditions that were presented on the line (I-Line/G-Line) compared to those on the pen (I-Pen/G-Pen), the performances of the two locations were nearly identical. The Orb-Line was most preferred and also present on the line, however, the reasons for preferring this visualization was primarily due to its non-user-based nature. Hence, more research will have to explore whether there are other effects to be found when presenting instructions at either of the locations.

The current study is also limited to following a straight line that does not change and does not include any curvatures or other variations. Within the design of our research, we considered adding curvature to the lines to study these effects. However, there were still too many gaps that first had to be answered in terms of speed acquisition for straight lines, which we tried to cover here. Future work should look at more variations and curvatures within the lines to study how the effects of the visualizations scale for other line types.

7 Conclusion

Our study contributes to the design of effective AR guidance visualizations for line tracing tasks. We identified three main aspects in our design: visualization location, user adaptation, and continuity. The results indicated that the demonstration on the line was the most effective visualization for motion-based guidance in terms of speed accuracy and the distance the speed was acquired at, at the cost of a larger speed overshoot at the start of the line tracing task. It is also important to consider the target speed to acquire since visualizations achieve different results for each target speed individually. These findings have important implications for the design of AR guidance for motor skill activities. By identifying the most effective visualizations for guiding users through a line tracing task, we can apply these insights to other motor skill activities that involve similar movements. This can help improve the design and implementation of AR guidance to assist users in learning and performing motor skill tasks more quickly and accurately while minimizing the risk of errors.

Acknowledgments

We would like to express our gratitude to Flanders Innovation and Entrepeneurship (VLAIO) in view of the DWI4Manufacturing project that kindly funded part of the research at the Expertise Centre for Digital Media. We also thank the participants for the time and effort they spent to participate in the user study.

References

[1] Ananya Ipsita, Levi Erickson, Yangzi Dong, Joey Huang, Alexa K Bushinski, Sraven Saradhi, Ana M Villanueva, Kylie A Peppler, Thomas S Redick, and Karthik Ramani. Towards modeling of virtual reality welding simulators to promote accessible and scalable training. doi: 10.1145/3491102.3517696

[2] J. Brooke. Sus: A 'quick and dirty' usability scale. 1996. doi: 10.1201/9781498710411-35
[3] J. Buchner, K. Buntins, and M. Kerres. The impact of augmented reality on cognitive load and performance: A systematic review. *Journal of Computer Assisted Learning*, 2021. doi: 10.1111/jcal.12617
[4] K.-E. Chang, J. Zhang, Y. S. Huang, Tzu Chien Liu, Tzu Chien Liu, T.-C. Liu, and Y.-T. Sung. Applying augmented reality in physical education on motor skills learning. 28(6):685–697. doi: 10.1080/10494820.2019.1636073
[5] M. Dastan, A. Uva, and M. Fiorentino. Gestalt driven augmented collimator widget for precise 5 dof dental drill tool positioning in 3d space. *International Symposium on Mixed and Augmented Reality*, 2022. doi: 10.1109/ismar55827.2022.00033
[6] S. Eom, D. Sykes, S. Rahimpour, and M. Gorlatova. Neurolens: Augmented reality-based contextual guidance through surgical tool tracking in neurosurgery. *International Symposium on Mixed and Augmented Reality*, 2022. doi: 10.1109/ismar55827.2022.00051
[7] M. S. Erden and A. Billard. Robotic assistance by impedance compensation for hand movements while manual welding. *IEEE Transactions on Systems, Man, and Cybernetics*, 2016. doi: 10.1109/tcyb.2015.2478656
[8] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. *Advances in psychology*, 1988. doi: 10.1016/s0166-4115(08)62386-9
[9] B. Lauber and M. Keller. Improving motor performance: selected aspects of augmented feedback in exercise and health. 14(1):36–43. doi: 10.1080/17461391.2012.725104
[10] J.-S. Liu, C. Elvezio, B. Tversky, and S. Feiner. Using multi-level precueing to improve performance in path-following tasks in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 2021. doi: 10.1109/tvcg.2021.3106476
[11] Z. Liu and J. Heer. The effects of interactive latency on exploratory visual analysis. *IEEE Transactions on Visualization and Computer Graphics*, 2014. doi: 10.1109/tvcg.2014.2346452
[12] A. Maeder. Human understanding limits in visualization. *International Journal of Pattern Recognition and Artificial Intelligence*, 1997. doi: 10.1142/s021800149700010x
[13] A. Nomoto, Y. Ban, T. Narumi, T. Tanikawa, and M. Hirose. Supporting precise manual-handling task using visuo-haptic interaction. *International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems*, 2016. doi: 10.1145/2875194.2875216
[14] L. Peternel, O. Sigaud, and J. Babič. Unifying speed-accuracy trade-off and cost-benefit trade-off in human reaching movements. *Frontiers in Human Neuroscience*, 2017. doi: 10.3389/fnhum.2017.00615
[15] R. Sodhi, H. Benko, and A. D. Wilson. Lightguide: projected visualizations for hand movement guidance. *International Conference on Human Factors in Computing Systems*, 2012. doi: 10.1145/2207676.2207702
[16] S. J. Thorpe, D. Fize, D. Fize, and C. Marlot. Speed of processing in the human visual system. *Nature*, 1996. doi: 10.1038/381520a0
[17] E. R. Tufte. Beautiful evidence. 2006.
[18] B. Volmer, J. Baumeister, S. V. Itzstein, I. Bornkessel-Schlesewsky, M. Schlesewsky, M. Billinghurst, and B. H. Thomas. A comparison of predictive spatial augmented reality cues for procedural tasks. *IEEE Transactions on Visualization and Computer Graphics*, 2018. doi: 10.1109/tvcg.2018.2868587
[19] S. J. Whitney and A. Stephens. Use of simulation to improve the effectiveness of army welding training.
[20] E. Wierinck, V. Puttemans, S. P. Swinnen, and D. van Steenberghe. Effect of augmented visual feedback from a virtual reality simulation system on manual dexterity training. 9(1):10–16. doi: 10.1111/j.1600-0579.2004.00351.x
[21] U. Yang and G. J. Kim. Implementation and evaluation of just follow me: an immersive, vr-based, motion-training system. *Presence: Teleoperators Virtual Environments*, 2002. doi: 10.1162/105474602317473240
[22] X. Yi, X. Wang, J. Li, and H. Li. Examining the fine motor control ability of linear hand movement in virtual reality. *IEEE Conference on Virtual Reality and 3D User Interfaces*, 2023. doi: 10.1109/vr55154.2023.00058
[23] X. Yu, K. Angerbauer, P. Mohr, D. Kalkofen, and M. Sedlmair. Perspective matters: Design implications for motion guidance in mixed reality. *International Symposium on Mixed and Augmented Reality*, 2020. doi: 10.1109/ismar50242.2020.00085
[24] Q. Zhou, A. Irlitti, D. Yu, J. Goncalves, and E. Velloso. Movement guidance using a mixed reality mirror. *Conference on Designing Interactive Systems*, 2022. doi: 10.1145/3532106.3533466
[25] L. Zimmerli, A. Duschau-Wicke, A. Mayr, R. Riener, L. Lünenburger, and L. Lünenburger. Virtual reality and gait rehabilitation augmented feedback for the lokomat. pp. 150–153. doi: 10.1109/icvr.2009.5174223