Understanding Perspectives for Single- and Multi-Limb Movement Guidance in Virtual 3D Environments

Hesham Elsayed
TU Darmstadt
Darmstadt, Germany
hesham.elsayed@tu-darmstadt.de

Kenneth Kartono
TU Darmstadt
Darmstadt, Germany
kenkart8@gmail.com

Dominik Schön
TU Darmstadt
Darmstadt, Germany
schoen@tk.tu-darmstadt.de

Martin Schmitz
Saarland University
Saarbrücken, Germany
mschmitz@cs.uni-saarland.de

Max Mühlhäuser
TU Darmstadt
Darmstadt, Germany
max@informatik.tu-darmstadt.de

Martin Weigel
Technische Hochschule Mittelhessen
Gießen, Germany
martin.weigel@mni.thm.de

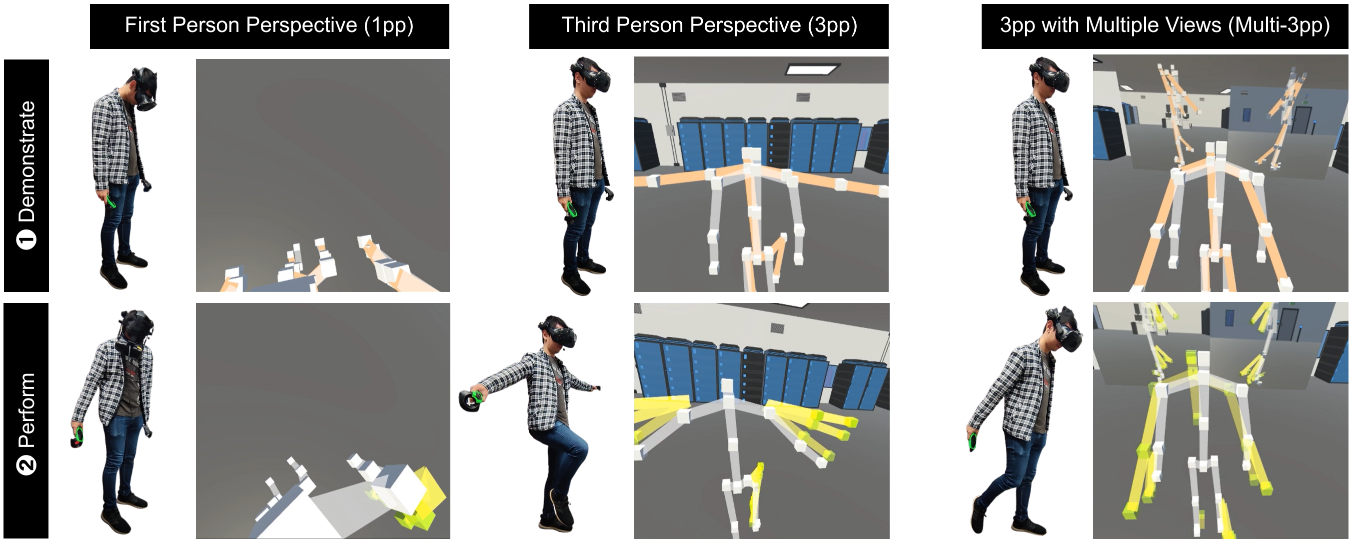

Figure 1: We evaluate three perspectives for motion guidance in VR: first-person, third-person, and multi third-person. The movement was first demonstrated by the system, before the users replicated the movement while seeing the steps of the motion.

ABSTRACT

Movement guidance in virtual reality has many applications ranging from physical therapy, assistive systems to sport learning. These movements range from simple single-limb to complex multi-limb movements. While VR supports many perspectives – e.g., first person and third person – it remains unclear how accurate these perspectives communicate different movements. In a user study (N=18), we investigated the influence of perspective, feedback, and movement properties on the accuracy of movement guidance. Participants had on average an angle error of 6.2° for single arm movements, 7.4° for synchronous two arm movements, and 10.3° for synchronous two arm and leg movements. Furthermore, the results show that the two variants of third-person perspectives outperform a first-person perspective for movement guidance (19.9% and 24.3% reduction in angle errors). Qualitative feedback confirms the quantitative data and shows users have a clear preference for third-person perspectives. Through our findings we provide guidance for designers and developers of future VR movement guidance systems.

CCS CONCEPTS

• **Human-centered computing** → **Interaction techniques**; **Virtual reality**; **Empirical studies in HCI**.

KEYWORDS

Movement guidance, body visualization, virtual reality

ACM Reference Format:
Hesham Elsayed, Kenneth Kartono, Dominik Schön, Martin Schmitz, Max Mühlhäuser, and Martin Weigel. 2022. Understanding Perspectives for Single- and Multi-Limb Movement Guidance in Virtual 3D Environments. In *28th ACM Symposium on Virtual Reality Software and Technology (VRST*

VRST '22, November 29-December 1, 2022, Tsukuba, Japan

ACM ISBN 978-1-4503-9889-3/22/11…$15.00
https://doi.org/10.1145/3562939.3565635

'22), November 29-December 1, 2022, Tsukuba, Japan. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3562939.3565635

1 INTRODUCTION

Movement guidance plays a vital role in many domains, ranging from sports training or physical therapy to dancing support. Traditionally, users train with a coach, who observes their movements and offers guidance in the form of visual (e.g., demonstration of a movement), auditory (e.g., instructions for moving certain body parts) and tactile (e.g., physical guidance through a movement) information. While this approach has proven to be effective [15], it depends on the availability of the coach and the user at the same time and place, and is limited by the attention span of coach/user and high costs. Moreover, during the recent pandemic's time, traditional training with a coach becomes even more unlikely.

To overcome these limitations, research has proposed a variety of approaches. Among the earliest, is the use of video tutorials for movement guidance [4]. However, this approach is very limited as it is hard to accurately decode movement information from a prerecorded video, and the user receives no feedback on their performance.

With the advent of virtual reality (VR) and low cost sensing solutions, new interaction techniques became possible, e.g. the ability to show the user movements from different perspectives. Research has shown that VR is a more effective medium for movement instruction compared to video based approaches [12]. Furthermore, for posture guidance [12] and path guidance [21], a first person perspective was shown to outperform other perspectives. However, posture and path guidance overlook the time dimension, users can look around frequently and take their time while performing the movements without adversely affecting the accuracy. In contrast, most real-world movements are performed with time playing an important role. Therefore, it remains unclear how users perform using different perspectives for movement guidance under time constraints.

The primary research question we investigate in this paper is how the perspective influences the accuracy of timed movements in virtual 3d environments (see Figure 1). Therefore, we conducted a controlled user study with 18 participants. We varied the perspective (1pp, 3pp, and Multi-3pp) and movement complexity (one arm, two arms, and two arms + leg) to understand their influence on different movements. We also varied movement direction (backward, forward, and sideways) and speed (fast and slow) to cover a large set of typical movements. Each condition was performed without real-time feedback on the participants performance, with visual or with haptic feedback.

Our results show that for movement guidance, a third person perspective with multiple views outperforms a first person perspective (24% decrease in average joint angle error). Furthermore, by increasing the body parts involved in the movement (i.e., the movement complexity), the ability of users to replicate the movement correctly decreases. The angular error increases from 6.2° for single arm movements to 7.4° for synchronous two arm movements and to 10.3° for synchronous two arm and leg movements. We further collected qualitative feedback through a survey. Users found VR to be a viable alternative for movement guidance and expressed clear preferences to using a third person perspective over a first person perspective.

In summary, the main contributions of this paper are:

(1) Findings from a controlled user study comparing three perspectives for single- and multi-limb movements in VR.
(2) Qualitative results from a survey investigating subjective preferences for VR movement guidance systems.
(3) A set of design recommendations based on our findings to help designers of VR movement guidance systems.

2 RELATED WORK

To contextualize our research and contributions, we describe the current state of the field in the following.

2.1 Perspectives in Virtual 3D Environments

Prior work has mainly used two types of perspectives for guidance in virtual reality: first-person and third-person perspective.

2.1.1 First-Person Perspective. The first-person perspective is the same as our real-world view of our bodies. Guidance in this case consists of visual cues superimposed on our view of our bodies. Although this perspective has been shown to be more effective than other perspectives for posture guidance [12], it also leads to constant head rotation in order to perceive guidance cues [21]. We hypothesize that for movement guidance, these constant head rotations would lead to a decreased accuracy as users are under time-constraints while performing the movement.

2.1.2 Third-Person Perspective. Commonly used in games, the third-person perspective shows an out-of-the-body view of the user. Two main types of third-person perspective were studied for guidance applications: mirror perspective (e.g., YouMove [2] and Physio @Home [17]) and a from behind view of a person as commonly found in games (e.g., the work by Yu et al. [21]). In our pilot tests, we found that the distance of visualization from the user influences the quality of guidance. The farther away a visualization was, the harder it was to perceive the movement. We therefore decided to use a third-person perspective from behind and above the user as it enabled placing the visualization close to the user. A mirror perspective would have to be placed farther away as the range of motion to the front is greater than to the back in the movements investigated. A view from behind also enabled executing the movements without mirroring them. We further used a skeleton visualization to reduce occlusions of the body that can make forward movements not visible.

2.2 Movement Guidance

Prior work on movement guidance can be grouped into (1) posture guidance, (2) path guidance, and (3) movement guidance.

2.2.1 Posture Guidance. Guidance of key frames in body movements is an important aspect of movement guidance and hence has been the subject of several research papers. OneBody [12] investigated the use of VR in comparison to video and Skype for remote posture guidance. Findings showed that using the first person perspective users could imitate target postures more accurately compared to Skype and prerecorded video. However, users also

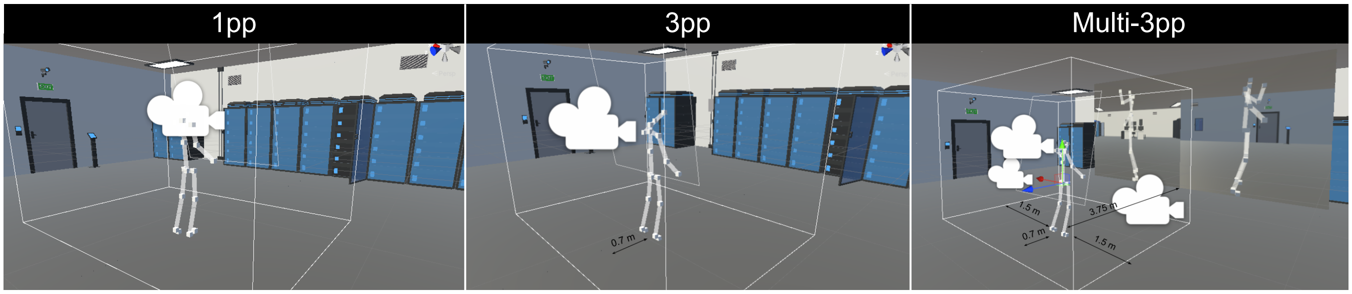

Figure 2: Illustration of the camera positions used by the different perspectives evaluated in our user study.

required the longest time to complete postures using a first person perspective. YouMove [2] was a system for posture- and movement guidance. Using an augmented mirror, users could see the target postures overlaid over their reflection. An evaluation showed that YouMove improved short-term retention compared to video demonstration. CameraReady [9] evaluated the use of different displays and visualizations for posture guidance. Larger displays led to lower errors and a 3d body visualization was rated to be more usable than a skeleton visualization.

2.2.2 Path Guidance. A step closer to movement guidance, path guidance focuses on guiding users along a predefined path in three-dimensional space. The main difference to movement guidance is that there are no time constraints on the movement. LightGuide [16] introduced the use of projected visualizations on the body for hand guidance. Results showed that users are 85% more accurate using these projected visualizations compared to guidance from animated videos. In a series of user studies, Yu et al. [21] investigated the effect of different perspectives on path guidance in VR. Findings showed that a first person perspective outperforms other perspectives. However, a third person perspective led to significantly lower values for head rotation, indicating that with first person perspective, users had to keep moving their heads left and right constantly to perceive the path accurately. OctoPocus3D [6] investigated feedforward and feedback for executing gestures in three-dimensional space. Findings showed that concurrent feedback is useful at the beginning, but as users execute the gestures more frequently it becomes unnecessary.

2.2.3 Movement Guidance. EGuide [7] investigated visual appearances and guidance techniques for mid-air arm movements. Findings showed that for continuous guidance, a realistic arm model resulted in higher accuracy compared to an abstract arm model. Hülsmann et al. [13] investigated showing users their own bodies, as well as showing users a superimposed body of the teacher from the front and side while performing squats. Results showed that users performed better with a superimposed visualization of skilled performance and that different views lead to different kinds of improvement. de Kok et al. [5] developed a closed loop system for multimodal feedback while learning movements. Movements of users are automatically evaluated by the system and corrective instructions are given to the users in real-time. Han et al. [11] introduced AR-Arm, a movement guidance system for upper limb motions from the first-person perspective. Although all these approaches have demonstrated movement guidance in VR, an evaluation of the influence of different perspective and movement characteristics is still missing in the literature. In our work, we systematically evaluate the affect of varying the perspective used as well as the movement on the accuracy of VR movement guidance.

3 CONCEPT

In the following, we describe and motivate the used concepts for the evaluated movement guidance.

3.1 Perspectives

This paper evaluates three perspectives (see Figure 1) and compares their influence on single- and multi-limb movement guidance.

3.1.1 First-Person Perspective (1pp). First-person shows the user's own body as a stick figure. The advantage of this perspective is that the user does not need to translate the movement, since they are already shown in the correct location. However, checking the accuracy of multiple limbs requires head movement.

3.1.2 Third-Person Perspective (3pp). Third-person shows a stick figure 0.7 m in front of the user (see Figure 2b). This distance was chosen to ensure the whole stick figure is in view without requiring head movements. Hence, we expect this perspective to perform better for multi-limb movements, where the user can not view all body parts. In contrast to a human trainer, where the trainer movements are mirrored to allow for eye contact between student and trainer, we decided to have the stick figure facing in the same direction as the user to ease the interpretation of the movements.

3.1.3 Third-Person Perspective with Multiple Views (Multi-3pp). Multi third-person is similar to the third-person condition, but shows two additional third-person views 3.75 m behind the third person view. The two additional perspectives show the user from both sides to give additional information about movements (see Figure 2c). We expect these views to help with forward and backward movements that are difficult to interpret in a third-person perspective.

Although 1pp and 3pp could be combined into a single perspective (similar to the multiple views in Multi-3pp), we make the distinction between 1pp and 3pp to focus on the mere influence of each perspective on its own. An interesting direction for future work can be investigating the combination of 1pp and 3pp.

3.2 Phases of Movement Training

The movement training consists of two phases: *demonstrate* and *perform*. The phases are inspired by interactions with human trainers (e.g., in a fitness course).

3.2.1 Demonstrate. The first phase shows the movement to the user. The movement is performed with the correct timing to help the user to understand the speed of the movements. Although this phase is similar to a demonstration by a human trainer, it differs in the perspective. Instead of viewing the mirrored movement on another human, the system uses the capabilities of a virtual environment to show the movement either as an overlay on the user, as a non-mirrored third-person or from multiple perspectives.

3.2.2 Perform. In the second phase the user performs the movement without demonstration. However, to guide the user's movements there are visible key frames along the movement path (see Figure 1). The guides are automatically extracted from the movement to ensure even spacing along the movement path. These visual guides disappear, when the user's body moves close enough (5 cm) to the position. In addition, we evaluated two types of feedback during this phase.

3.3 Feedback

We evaluate movement guidance without and with two feedback modes in the perform phase. These modes were added to better understand the influence of visual and haptic feedback during multi-limb movements, since they might ease the adjustment of multiple body parts and hence lead to more accurate movements.

3.3.1 None. The baseline condition shows the the keyframes without additional feedback based on the user's movement.

3.3.2 Color Feedback. The color condition shows keyframes and *changes the color* of the VR stick figure's body parts depending on how close these body parts are to their optimal path. If the distance between a body joint and the same joint in the optimal movement exceeds 15 cm, the joint is colored red.

3.3.3 Haptic Feedback. The haptic condition shows key frames and gives *vibrotactile feedback* on each of the body parts depending on how close they are to their optimal path. The haptic stimulus was calibrated for each user and we used a linear mapping between the euclidean distance to the optimal path (distances beyond 15 cm are mapped to 100% and distances under 15 cm stopped the vibration).

4 USER STUDY

Our user study investigates the influence of different perspectives, movement types, and feedback on the accuracy of movement guidance in VR. In particular, the quantitative part of our user study aims to answer the following hypotheses:

H1 Users are more accurate with 1pp for movements involving one arm.
Motivation: prior work on posture guidance showed that for upper limb postures, 1pp was more accurate than 3pp. However, since our study investigates movement guidance, we hypothesize that one arm movements will not require frequent head rotations and hence will be more accurate using 1pp.

H2 Users are more accurate with 1pp for forward movements.
Motivation: we based this hypothesis on the fact that users were more accurate with 1pp for visible postures and paths from prior work [12, 21].

H3 Users are more accurate with 3pp and Multi-3pp for movements requiring multiple body parts.
Motivation: we hypothesize that due to frequent head rotations necessary to perceive multi-limb movements accurately with 1pp, 3pp and Multi-3pp will be more accurate.

H4 Users are more accurate with Multi-3pp than 3pp.
Motivation: we hypothesize that the addition of side views in Multi-3pp will enable users to see errors in their movements more easily, as was shown in prior work [13] for other perspectives.

H5 Users perform movements requiring fewer body parts more accurately.
Motivation: we hypothesize that multi-limb movements require higher coordination efforts from users and hence result in higher movement errors.

H6 Users are more accurate with haptic feedback than color feedback and no feedback.
Motivation: haptic feedback can be instantaneously localized and does not require visual attention to the body part. Hence, we hypothesize that haptic feedback will be more effective than color and no-feedback in correcting errors.

H7 Users perform slow movements more accurately than fast movements.
Motivation: we hypothesize that slow movements can be more easily replicated than faster movements.

4.1 Participants

We recruited 18 (13 male, 5 female) participants aged between 21 and 27 years old ($\mu = 23.89$, $\sigma = 1.94$). Three of our participants had previous experiences with VR. One participant had previously explored a museum in VR, and the remaining two used VR for gaming purposes. None of our participants had previous experience with movement guidance in VR. Participation in our experiment was voluntary, and no compensation was offered.

4.2 User Study Design

In our user study, we varied the *perspective* (1pp, 3pp, and Multi-3pp), *movement complexity* (one arm, two arms, and two arms + leg), *movement direction* (forward, sideways, and backward), *movement speed* (slow and fast), and *feedback* (none, haptic, and color). This resulted in 162 ($3 \times 3 \times 3 \times 2 \times 3$) conditions. We used a balanced Latin-square to counterbalance the variable *perspective*. The order of the remaining variables was randomized. In total, we collected from each participant 162 movements.

4.3 Procedure

After obtaining informed consent from the participants, we collected their demographic data. Then, we explained the task and provided a brief overview of the procedure. The task was to replicate a movement after observing it in VR.

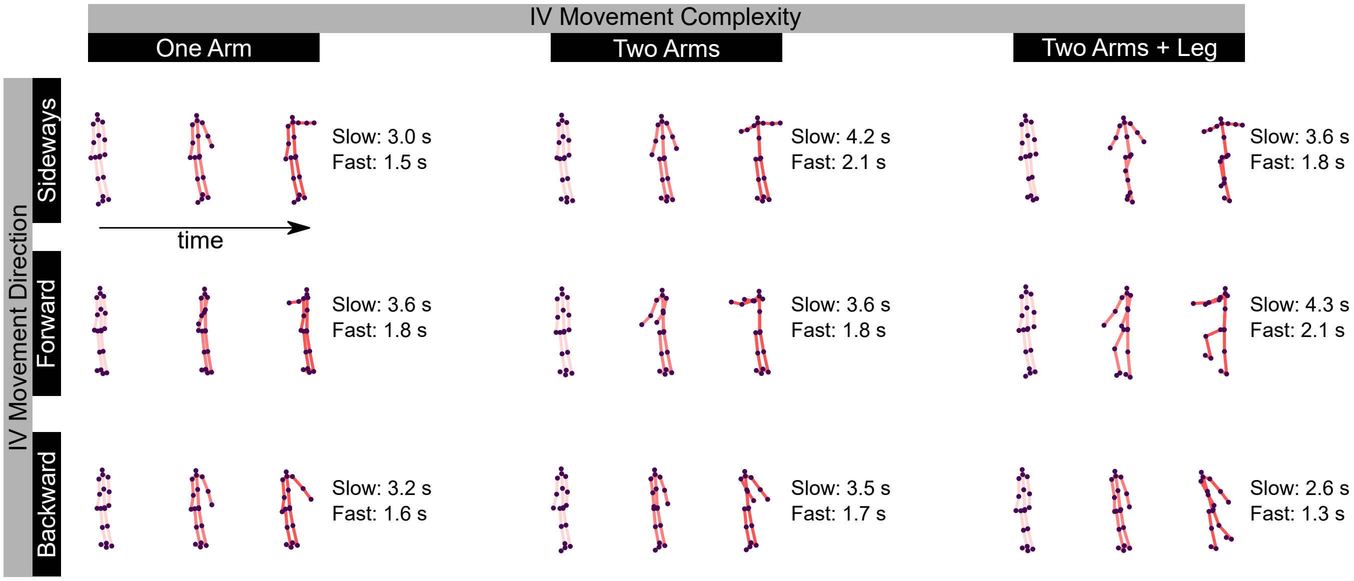

Figure 3: The nine different movements used in our evaluation. The levels of the two independent variables *Movement Complexity* and *Movement Direction* are illustrated.

Every trial started with the participant standing in a neutral pose with hands to the side. Upon pressing the trigger button of the Vive controller, a movement is visualized for the participant to observe. Pressing the trigger button again signals that the participant is ready to start. A guiding visualization is displayed and the participant can start moving. Upon completing a movement, the participant presses the trigger button once again to indicate that the movement is finished. When to press the trigger button was explained to the participants at the beginning, and 2-3 test movements were performed to get the participants familiar with our system that were not recorded. Participants were instructed to replicate the movements in the same speed in which they are displayed.

Upon completing all movements in a certain *perspective*, participants took a small break of 5–10 minutes. After completing all movements, we collected qualitative feedback by asking participants to fill out a survey with the following questions:

S1 What is your opinion on using a VR system to learn new movements?
S2 Which would you prefer: a VR system, a TV application or a real class for learning new movements? Why?
S3 Which perspective did you like the most in VR? Why?
S4 Which feedback did you like the most? Why?
S5 Did you find any aspects frustrating while using VR for motion guidance? Which?
S6 Are there further features you would like to see in a VR movement guidance application? Which?

The total duration of the experiment was approx. 60 minutes.

4.4 Apparatus

We conducted the experiment on a i7 dual core 3.6 GHz, 16 GB RAM desktop PC with a NVIDIA GeForce GTX 970 graphics card. We used an HTC VIVE headset and Microsoft Kinect v2. Although not as accurate and precise as marker-based systems, the Microsoft Kinect v2 is accurate and depending on the joint, moderately precise [20]. The virtual environment was running on the same desktop computer and updated tracking information at 60 Hz.

Haptic feedback was generated using the built-in linear resonant actuators (LRAs) in the HTC Vive controllers. In addition to the controllers, haptic feedback was required for stimulation of the legs. We used two EAI C2 [3] linear actuators attached to the ankles. The actuators were set to 200 Hz and vibrated at full intensity with a maximum peak to peak displacement of 0.8 mm. They were placed directly under the outer side of the ankle (i.e., under the Lateral Malleolus bone), as vibrations on the bone were sometimes perceived as uncomfortable by the participants.

4.5 Dependent Variables

We recorded the joint angle errors of the participants while performing movements. The joint angle error was computed frame by frame for all joints involved in the movements (i.e., the shoulders, elbows, hips, and knees) and averaged over joints and frames to produce a single value per movement. The following formula was used to calculate the joint angle denoted by θ:

$$\theta = \arccos\left(\frac{a \cdot b}{|a||b|}\right)$$

Where a and b are the vectors connecting the three joints, e.g., for the shoulder a is the vector connecting the shoulder to the spine and b is the vector connecting the shoulder to the elbow. We recorded the 3d positional accuracy (euclidean distance), but decided against using it for further analysis, as it was sensitive to participants' body sizes and full body translations, e.g., a participant stepping forward while performing a movement.

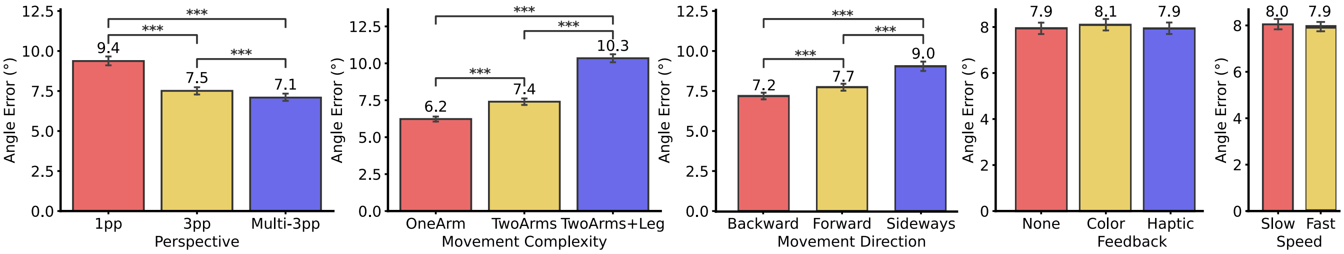

Figure 4: Joint angle errors for all independent variables. Error bars are the 95% confidence intervals. All significant effects are shown ($^{*} \leq 0.05$, $^{**} \leq 0.01$, and $^{***} \leq 0.001$).

4.6 Data Analysis

We tested the data for normality using the Shapiro-Wilk test and found significant deviations. We therefore decided to use the Aligned Rank Transform (ART) [19] procedure to process our data. We then performed an ANOVA to compute the F-score and p-value of main and interaction effects as suggested by Wobbrock et al. [19]. As the ART procedure can inflate Type I errors for post-hoc pairwise comparisons, we used ART-C [8] for post-hoc testing with Bonferroni corrections.

5 QUANTITATIVE RESULTS

We analyse the angle errors from our study and discuss the main (see Figure 4) and interaction effects (see Figure 5).

5.1 Main Effects

The analysis showed a significant ($F_{2,2734}$ = 192.49, p < .001) main effect of the variable *perspective* on the joint angle error. We found that 1pp (μ = 9.37 °, σ = 4.62 °) resulted in the highest joint angle errors, followed by 3pp (μ = 7.51 °, σ = 3.51 °), and Multi-3pp (μ = 7.09 °, σ = 3.51 °). Post-hoc tests confirmed significant differences between 1pp and 3pp (p < .001), 1pp and Multi-3pp (p < .001), and 3pp and Multi-3pp (p < .001).

We found a statistically significant ($F_{2,2734}$ = 737.45, p <.001) main effect of the variable *movement complexity* on the joint angle error of participants. Movements involving the use of one arm (μ = 6.23 °, σ = 2.90 °) resulted in the lowest joint angle errors, followed by two arms (μ = 7.40 °, σ = 3.47 °), and finally two arms and a leg (μ = 10.35 °, σ = 4.42 °). Post-hoc tests confirmed significant differences between one arm and two arm (p < .001) movements, one arm and two arms plus leg (p < .001), and between two arms and two arms plus leg (p < .001).

Further, we found a significant ($F_{2,2734}$ = 147.43, p <.001) main effect of the variable *movement direction* on the joint angle errors of participants. Backward (μ = 7.19 °, σ = 3.58 °) movements resulted in lowest joint angle errors, followed by forward (μ = 7.74 °, σ = 3.59 °) and sideway (μ = 9.05 °, σ = 4.63 °) movements. Post-hoc tests confirmed all pair-wise differences as significant (p < .001).

We found no significant main effects for the variables *feedback* ($F_{2,2734}$ = 1.18, p >.05) and *speed* ($F_{1,2734}$ = 0.36, p >.05). Figure 4 summarizes the results.

5.2 Interaction Effects

Figure 5 displays the two-way interactions.

5.2.1 *Movement Direction * Perspective.*

We found a significant ($F_{4,2734}$ = 7.60, p <.001) interaction effect between the variables *perspective* and *movement direction*. For 1pp, backward and forward movements were comparable, whereas for 3pp and Multi-3pp, backward movements were significantly more accurate than forward movements.

5.2.2 *Movement Direction * Movement Complexity.*

We found a significant interaction effect ($F_{4,2734}$ = 37.53, p <.001) between the variables *movement complexity* and *movement direction*. Backward movements were significantly more accurate than forward movements only for movements with two arms and a leg. For movements with one arm and two arms, no significant difference was found between backward and forward directions.

5.2.3 *Movement Direction * Feedback.*

We did not find a significant ($F_{4,2734}$ = 0.50, p >.05) interaction effect between the variables *feedback* and *movement direction*.

5.2.4 *Feedback * Perspective.*

Our analysis did not reveal a significant ($F_{4,2734}$ = 0.42, p >.05) interaction effect between the variables *perspective* and *feedback*.

5.2.5 *Movement Complexity * Perspective.*

Our analysis did not reveal a significant ($F_{4,2734}$ = 1.17, p >.05) interaction effect between the variables *perspective* and *movement complexity*.

5.2.6 *Movement Complexity * Feedback.*

We did not find a significant ($F_{4,2734}$ = 1.02, p >.05) interaction effect between the variables *feedback* and *movement complexity*.

5.2.7 *Speed * Perspective.*

We found a significant ($F_{2,2734}$ = 3.70, p <.05) interaction effect between the variables *perspective* and *speed*. For 1pp, fast movements had a higher joint angle error than slow movements. On the other hand, for 3pp and Multi-3pp, fast movements had a lower joint angle error. Post-hoc testing did not confirm these differences as significant.

5.2.8 *Speed * Movement Direction.*

Our analysis revealed a significant ($F_{2,2734}$ = 36.97, p <.001) interaction effect between the variables *movement direction* and *speed*. Fast movements were significantly more accurate than slow movements for backward and forward directions. For the movement direction sideways, slow movements were significantly more accurate than fast movements.

5.2.9 *Speed * Movement Complexity.*

We found a significant ($F_{2,2734}$ = 4.87, p <.01) interaction effect between the variables *movement*

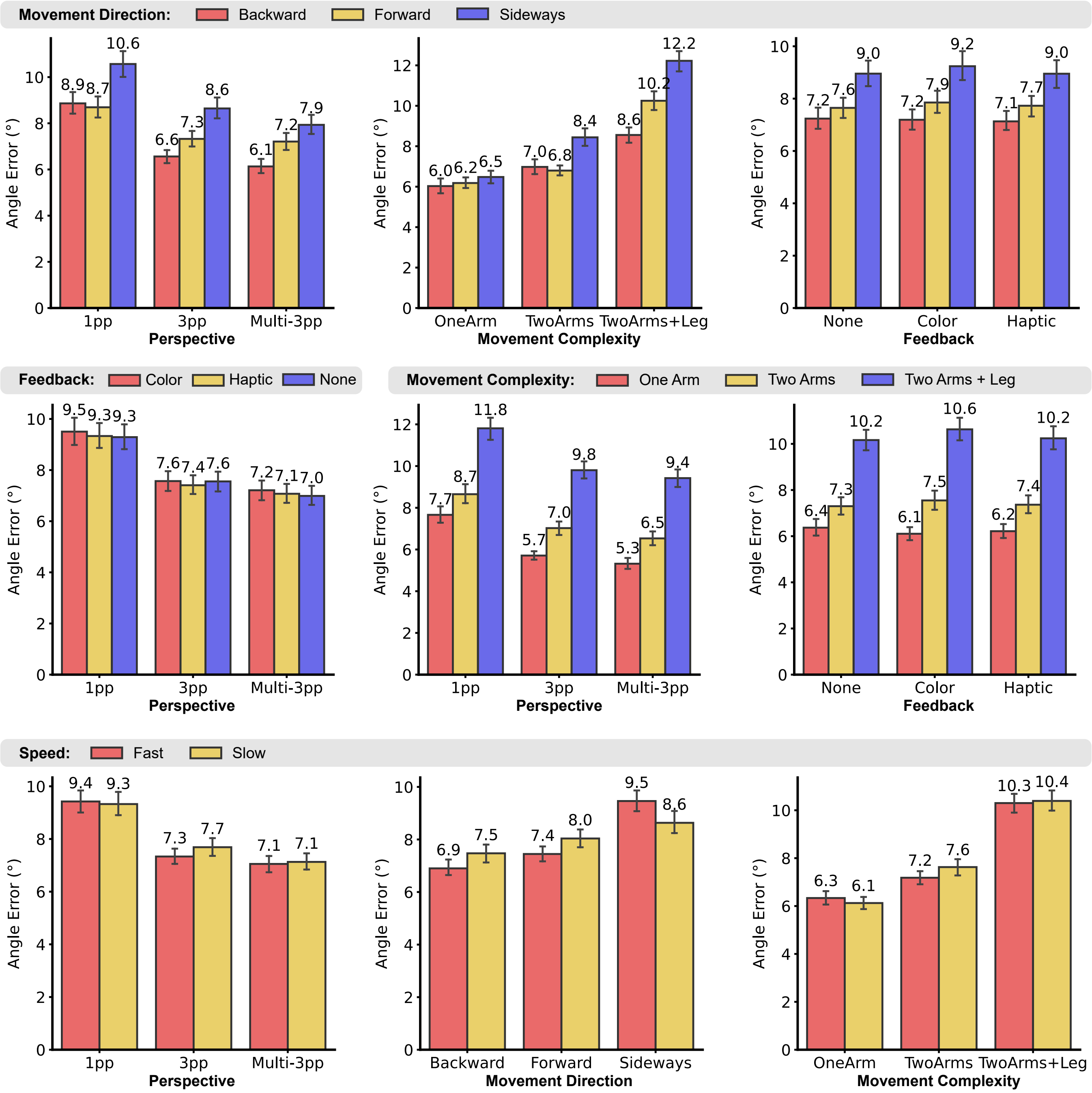

Figure 5: Joint angle errors for two-way interactions between the independent variables. Error bars are 95% confidence intervals.

complexity and *speed*. One arm slow movements were more accurate than fast movements, while two arm fast movements were more accurate than slow movements. Post-hoc tests did not confirm these differences as significant.

*5.2.10 Speed * Feedback.* We did not find a significant ($F_{2,2734}$ = 0.83, p >.05) interaction effect between the variables *feedback* and *speed*.

6 QUALITATIVE RESULTS

In the following, we detail on the results of our survey.

6.1 Opinions on Using VR to Learn Movements

The majority of participants (16) expressed a positive attitude towards using a VR system to learn new movements. They found it *"very interesting"* (P1, P4, P5, P10, P12), *"engaging and fun"* (P1, P10, P15), and it to be *"a good idea"* (P8, P9, P11, P13). Some participants expressed that the usefulness of the system depended on the perspective being used: *"pretty useful from the third person perspective"* (P7), *"it is kinda hard to watch the movement, especially in the first person point of view. With the third person's point of view it is easier to see the movement"* (P15), and that *"the third person perspective is easier to follow."* (P9). Two participants expressed negative attitudes towards learning new movements with a VR system. They found it *"quite tiring"* (P6) and it to be *"not very effective"* (P14).

6.2 Preferences: Real Class, TV, or VR

Nine participants expressed preferences for learning new movements in a real class. P1 thought that *"an expert would be monitoring my actions and provide a more accurate feedback"*. Similarly, P7 thought that a real class provides *"quick and easier instruction from a person"* and P8 expressed that with a real class it is *"easier to follow the movement and someone to correct the movement if I did it wrongly"*. The remaining participants all expressed similar thoughts, with the presence of an expert to support them being the main reason why they would prefer a real class. Seven participants chose a VR system for learning new movements. P3 appreciated the independence associated with VR: *"I can learn the movements myself"*. P9 appreciated the flexibility: *"with VR I can do it at home and whenever I like"*. The remaining participants expressed similar thoughts, with independence being the main reason why they prefer a VR system. Lastly, two participants expressed preferences towards a TV system. The main reason was that with a TV, a headset is not required: *"no need for a heavy headset to wear on the head"* (P10).

6.3 Preferences: Perspectives in VR

Ten participants preferred the use of Multi-3pp, while the remaining 8 participants all chose 3pp with no one choosing 1pp. The main reason users gave for not choosing 1pp was that 3pp and Multi-3pp were *"clearer"* (P1, P5, P13, P14) and allow users to *"observe the full movement"* (P4, P7), whereas 1pp caused *"neck pain"* (P1, P10) and required *"constant shifting from looking left to looking right"* (P1). Participants appreciated the ability to see themselves from *"different angles"* (P12) in Multi-3pp. Furthermore, P8 expressed *"if I am not sure from the 3rd person perspective, the other views makes it clear"*. On the other hand, P13 expressed that with 3pp *"no distraction compared to third person with multiple view"*. Similarly, P1 expressed that the multi-views can sometimes feel *"redundant"*, however, they are *"nice to have"*.

6.4 Preferences: Feedback

13 participants preferred haptic feedback, 3 participants chose color feedback, and 2 participants preferred no feedback. Arguments for haptic feedback included: *"enabled me to respond quickly"* (P6), *"I can feel it directly"* (P18), *"don't need to see it, I can just feel it"* (P16), and *"I can feel the feedback instantly"* (P10). P15 further expressed that the haptic feedback felt more *"fun...it feels more like playing games instead of learning a new movement"*. The main argument for color feedback was the ability to *"see the feedback directly"* (P12). Participants that preferred no feedback expressed that *"I did not understand the feedback and what was it telling me"* (P8) and *"other types of feedback are not very clear"* (P14).

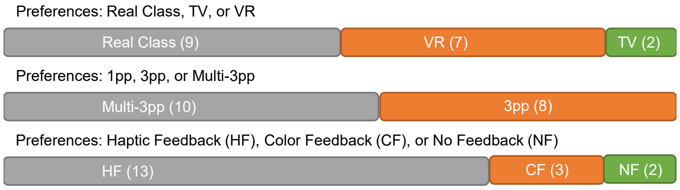

Figure 6: Users' qualitative preferences.

6.5 Frustrating Aspects

We asked participants if they found any aspects frustrating while using VR for movement guidance. Nine participants expressed that movement guidance with 1pp was frustrating. P18 expressed *"first person perspective is really tiring as I have to look around multiple times to see the movement"*. Similarly, P15 expressed *"the first person POV, it's hard to see the backside, i need to look around so much, pain to my neck"*. Three participants commented on the VR headset being *"quite heavy"* (P4, P16) and that it is *"not portable"* (P5).

6.6 Further Features in a VR Movement Guidance Application

Further features that participants wanted to have in a VR movement guidance application were: mirrored perspective as it *"might be more natural to the users, it is also a more common perspective when following an instructor in a real class."* (P1), better tracking, adding audio to the experience, being able to choose scenery, having the option to change the skeleton to look like a real instructor, and adding more movements, e.g., jumping, walking, and side steps.

7 DISCUSSION

In this section, we discuss quantitative and qualitative results of our user study. In general, we found that 3pp and Multi-3pp lead to significantly more accurate execution of movements compared to 1pp. This was also reflected in the qualitative results, where 10 participants chose Multi-3pp as their preferred perspective and the remaining participants choosing 3pp.

7.1 Quantitative Results

We hypothesized at the beginning in **H1** and **H2** that users would be more accurate using 1pp for one arm movements and forward movements, respectively. However, our results showed that for all levels of movement complexity (one arm, two arms, two arms and a leg) as well as all levels of movement direction (forward, backward, and sideways), 3pp and Multi-3pp were more accurate than 1pp. Therefore, we cannot support **H1** & **H2**. An observation that we could make explaining this result is that regardless of the movement complexity and the direction, users had to nevertheless look around to make sure that there is no visual information that they missed. This led to higher errors using 1pp, even for one-arm movements and movements forward.

In **H3**, we hypothesized that users would be more accurate for multi-limb movements with 3pp and Multi-3pp. Our results confirm this and hence we can accept **H3**. In **H4**, we hypothesized that Multi-3pp would lead to lower errors in comparison to 3pp. Our results confirm that Multi-3pp resulted in significantly lower errors in comparison to 3pp and 1pp, and hence we can accept **H4**. Our findings further show that multi-limb movements result in higher errors compared to single-limb movements, and therefore we can accept **H5**.

In **H6**, we hypothesized that haptic feedback would lead to lower errors in comparison to color and no-feedback. However, our results show that performance of participants was comparable over all feedback types and no significant differences were found. Therefore, we cannot support **H6**. A possible reason for the lack of improvement of haptic and color feedback in comparison to no-feedback was provided by our participants, where they stated that the feedback is useful to find out if a deviation from the target movement is being made, but it does not inform the user in which direction the error is being made. Hence, participants were unsure how to correct their movements after being notified by the different feedback types.

Regarding the speed of movement, we hypothesized that fast movements would be harder to execute than slow (**H7**). Our results showed that users executed fast and slow movements comparably, and hence we cannot support **H7**.

7.2 Subjective Preferences

In general, our participants expressed positive attitudes towards movement guidance in VR. However, 9 of our participants (50%) preferred a real class over movement guidance in VR, with the main reason being the presence of an expert that monitors and corrects movements. Therefore, future systems for movement guidance in VR should focus on supporting the interaction between an expert and the user, either using automated algorithms [14] for error correction or by supporting a two-way communication channel between expert and student in VR. This approach has the potential of combining the best of training in a real-class (presence of expert) and training in VR (independence).

In line with the quantitative results, all our participants expressed preference to either Multi-3pp (56%) or 3pp (44%). Hence, to improve the user experience of movement guidance systems in VR, a third-person perspective should be used.

Regarding the choice of feedback, haptic feedback was preferred by the majority of our participants (72%). Although the use of haptic feedback did not lead to a significant improvement in our quantitative results, it was preferred by participants over color and no-feedback, as it enabled them to respond quickly, did not require their visual attention, and felt more fun. To improve the user experience, future movement guidance systems should include haptic feedback and if possible encode the direction of correction required, e.g., using the push and pull metaphors [10], so that users are certain about what the feedback is communicating and the correction required.

8 DESIGN RECOMMENDATIONS

Based on our results, we derive four design recommendations for designers, developers, and researchers of motion guidance systems.

8.1 Third-Person Perspectives for Movements

Both third-person perspectives (3pp and Multi-3pp) performed significantly better than the first-person perspective (1pp) for movement guidance. Third-person perspectives lead to an angle error reduction 19.9% or 24.3% compared to the first-person perspective.

8.2 Additional Views increase Accuracy

We recommend adding multiple views for best accuracy, because Multi-3pp resulted in the least angle errors overall. The two additional views significantly reduced angle errors by 5.3% (-0.4°) compared to a third-person perspective without additional views.

8.3 Single-Limb Gestures for Precise Input

New interaction possibilities that use movement gestures as an input modality are constantly being introduced by researchers, e.g., for interaction with public displays [18] or mid-air gestures [1]. The study shows that a lower movement complexity leads to significantly more accurate movements. We recommend to use single-limb movement gestures whenever precise input is required (-15.8% compared to two arms and -39.8% compared to two arms and leg).

8.4 Improve User Experience with Haptics

Despite no significant effect on the movement errors, the qualitative feedback revealed 13 of 18 participants (72%) preferred haptic feedback over no feedback or color feedback. Therefore, we recommend adding haptic feedback to communicate errors to improve the user experience of motion guidance systems.

9 CONCLUSION

In this paper, we investigated different perspectives, movement properties, and feedback on VR movement guidance. Our work extends prior work on posture [2, 9, 12], path [6, 16, 21], and movement guidance [5, 7, 13]. Our findings show that for timed movements, a third-person perspective should be preferred as it increases the accuracy of the user while replicating movements. It was also qualitatively preferred by our users. Furthermore, our results showed that the addition of multiple views led to a significant accuracy increase and that single-limb movements were more accurately replicated in comparison to multi-limb movements. Based on our quantitative and qualitative findings, we derive a set of design recommendations for VR movement guidance systems.

ACKNOWLEDGMENTS

We would like to thank the reviewers for their feedback. This work has been funded by the German Federal Ministry of Education and Research (01IS17050).

REFERENCES

[1] Christopher Ackad, Andrew Clayphan, Martin Tomitsch, and Judy Kay. 2015. An In-the-Wild Study of Learning Mid-Air Gestures to Browse Hierarchical Information at a Large Interactive Public Display. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Osaka, Japan) *(UbiComp '15)*. Association for Computing Machinery, New York, NY, USA, 1227–1238. https://doi.org/10.1145/2750858.2807532

[2] Fraser Anderson, Tovi Grossman, Justin Matejka, and George Fitzmaurice. 2013. YouMove: Enhancing Movement Training with an Augmented Reality Mirror. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology* (St. Andrews, Scotland, United Kingdom) *(UIST '13)*. Association for

Computing Machinery, New York, NY, USA, 311–320. https://doi.org/10.1145/2501988.2502045

[3] Lorna Brown, Stephen Brewster, and Helen Purchase. 2006. Multidimensional tactons for non-visual information presentation in mobile devices. *ACM International Conference Proceeding Series* 159, 231–238. https://doi.org/10.1145/1152215.1152265

[4] Christopher Clarke, Doga Cavdir, Patrick Chiu, Laurent Denoue, and Don Kimber. 2020. *Reactive Video: Adaptive Video Playback Based on User Motion for Supporting Physical Activity*. Association for Computing Machinery, New York, NY, USA, 196–208. https://doi.org/10.1145/3379337.3415591

[5] Iwan de Kok, Julian Hough, Felix Hülsmann, Mario Botsch, David Schlangen, and Stefan Kopp. 2015. A Multimodal System for Real-Time Action Instruction in Motor Skill Learning. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction* (Seattle, Washington, USA) *(ICMI '15)*. Association for Computing Machinery, New York, NY, USA, 355–362. https://doi.org/10.1145/2818346.2820746

[6] William Delamare, Thomas Janssoone, Céline Coutrix, and Laurence Nigay. 2016. Designing 3D Gesture Guidance: Visual Feedback and Feedforward Design Options. 152–159. https://doi.org/10.1145/2909132.2909260

[7] Maximilian Dürr, Rebecca Weber, Ulrike Pfeil, and Harald Reiterer. 2020. EGuide: Investigating Different Visual Appearances and Guidance Techniques for Egocentric Guidance Visualizations. In *Proceedings of the Fourteenth International Conference on Tangible, Embedded, and Embodied Interaction* (Sydney NSW, Australia) *(TEI '20)*. Association for Computing Machinery, New York, NY, USA, 311–322. https://doi.org/10.1145/3374920.3374945

[8] Lisa A. Elkin, Matthew Kay, James J. Higgins, and Jacob O. Wobbrock. 2021. An Aligned Rank Transform Procedure for Multifactor Contrast Tests. In *The 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '21)*. Association for Computing Machinery, New York, NY, USA, 754–768. https://doi.org/10.1145/3472749.3474784

[9] Hesham Elsayed, Philipp Hoffmann, Sebastian Günther, Martin Schmitz, Martin Weigel, Max Mühlhäuser, and Florian Müller. 2021. CameraReady: Assessing the Influence of Display Types and Visualizations on Posture Guidance. In *Designing Interactive Systems Conference 2021* (Virtual Event, USA) *(DIS '21)*. Association for Computing Machinery, New York, NY, USA, 1046–1055. https://doi.org/10.1145/3461778.3462026

[10] Sebastian Günther, Florian Müller, Markus Funk, Jan Kirchner, Niloofar Dezfuli, and Max Mühlhäuser. 2018. TactileGlove: Assistive Spatial Guidance in 3D Space through Vibrotactile Navigation. In *Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference* (Corfu, Greece) *(PETRA '18)*. Association for Computing Machinery, New York, NY, USA, 273–280. https://doi.org/10.1145/3197768.3197785

[11] Ping-Hsuan Han, Kuan-Wen Chen, Chen-Hsin Hsieh, Yu-Jie Huang, and Yi-Ping Hung. 2016. AR-Arm: Augmented Visualization for Guiding Arm Movement in the First-Person Perspective. In *Proceedings of the 7th Augmented Human International Conference 2016* (Geneva, Switzerland) *(AH '16)*. Association for Computing Machinery, New York, NY, USA, Article 31, 4 pages. https://doi.org/10.1145/2875194.2875237

[12] Thuong N. Hoang, Martin Reinoso, Frank Vetere, and Egemen Tanin. 2016. Onebody: Remote Posture Guidance System Using First Person View in Virtual Environment. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction* (Gothenburg, Sweden) *(NordiCHI '16)*. Association for Computing Machinery, New York, NY, USA, Article 25, 10 pages. https://doi.org/10.1145/2971485.2971521

[13] Felix Hülsmann, Cornelia Frank, Irene Senna, Marc O. Ernst, Thomas Schack, and Mario Botsch. 2019. Superimposed Skilled Performance in a Virtual Mirror Improves Motor Performance and Cognitive Representation of a Full Body Motor Action. *Frontiers in Robotics and AI* 6 (2019). https://doi.org/10.3389/frobt.2019.00043

[14] Felix Hülsmann, Jan Göpfert, Barbara Hammer, Stefan Kopp, and Mario Botsch. 2018. Classification of motor errors to provide real-time feedback for sports coaching in virtual reality — A case study in squats and Tai Chi pushes. *Computers & Graphics* 76 (08 2018). https://doi.org/10.1016/j.cag.2018.08.003

[15] Charles Shea and Rebecca Lewthwaite. 2010. Motor skill learning and performance: A review of influential factors. *Medical education* 44 (01 2010), 75–84. https://doi.org/10.1111/j.1365-2923.2009.03421.x

[16] Rajinder Sodhi, Hrvoje Benko, and Andrew Wilson. 2012. LightGuide: Projected Visualizations for Hand Movement Guidance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 179–188. https://doi.org/10.1145/2207676.2207702

[17] Richard Tang, Xing-Dong Yang, Scott Bateman, Joaquim Jorge, and Anthony Tang. 2015. Physio@Home: Exploring Visual Guidance and Feedback Techniques for Physiotherapy Exercises. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 4123–4132. https://doi.org/10.1145/2702123.2702401

[18] Robert Walter, Gilles Bailly, and Jörg Müller. 2013. StrikeAPose: Revealing Mid-Air Gestures on Public Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 841–850. https://doi.org/10.1145/2470654.2470774

[19] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963

[20] Kirk Woolford. 2015. Defining Accuracy in the Use of Kinect v2 for Exercise Monitoring *(MOCO '15)*. Association for Computing Machinery, New York, NY, USA, 112–119. https://doi.org/10.1145/2790994.2791002

[21] Xingyao Yu, Katrin Angerbauer, Peter Mohr, Denis Kalkofen, and Michael Sedlmair. 2020. Perspective Matters: Design Implications for Motion Guidance in Mixed Reality. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 577–587. https://doi.org/10.1109/ISMAR50242.2020.00085