

Design Space of Visual Feedforward And Corrective Feedback in XR-Based Motion Guidance Systems

Xingyao Yu

Xingyao.Yu@visus.uni-stuttgart.de
University of Stuttgart
Stuttgart, Germany

Benjamin Lee

Benjamin.Lee@visus.uni-stuttgart.de
University of Stuttgart
Stuttgart, Germany

Michael Sedlmair

Michael.Sedlmair@visus.uni-stuttgart.de
University of Stuttgart
Stuttgart, Germany

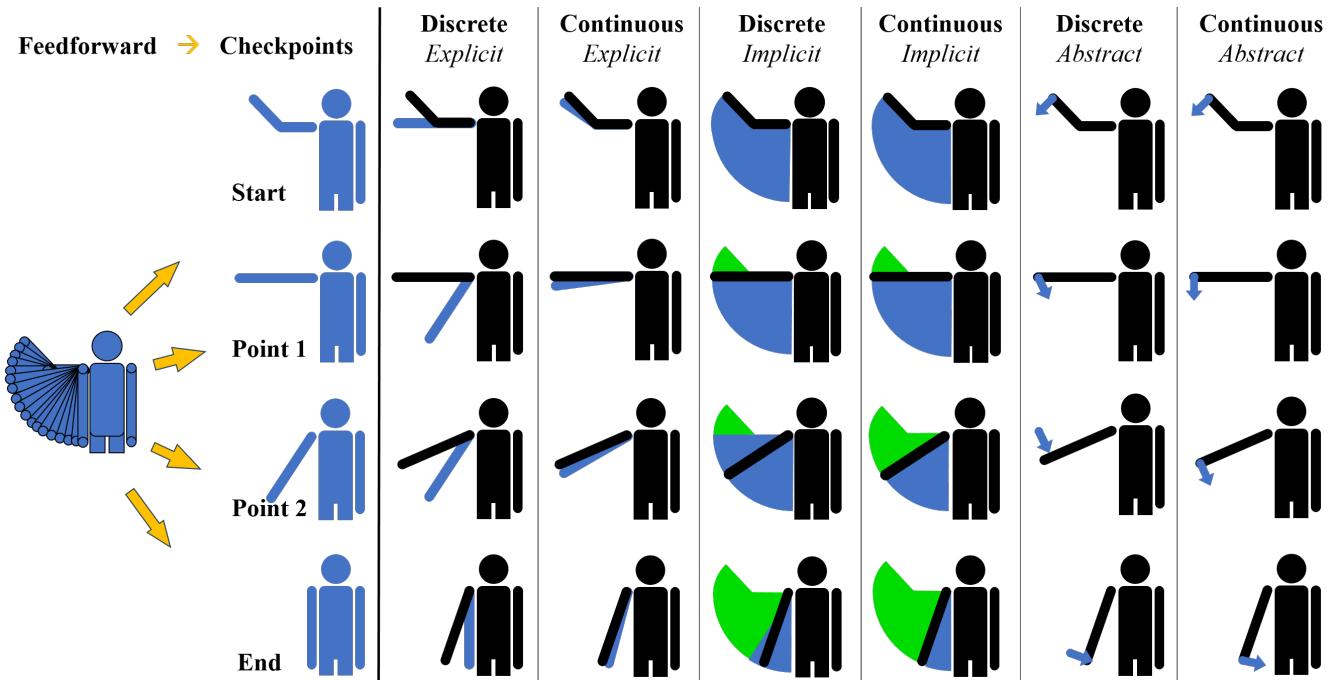


Figure 1: Schematic representations of motion feedforward. The leftmost column shows the complete motion sequence and its critical postures at four critical moments (*Start*, *Point1*, *Point2*, and *End*). The remaining columns show how the feedforward of this motion can be represented using different configurations of level of indirection (*explicit*, *implicit* and *abstract*) and interactive update strategy (*discrete*, *continuous* and *autonomous*). The visualizations in **blue** represent the feedforward; while the the black glyphs show the trainee's four moments during training from top to bottom (*Start*, *reach Point 1*, *between Point 1 and Point 2*, and *between Point 2 and End*); the **green** parts visualize the completed parts of a motion when using **Implicit** guidance (column 4 and 5). **Discrete** guidance (column 2, 4 and 6) presents cues for the next critical posture, which updates only when that posture is satisfied. **Continuous** guidance (column 3, 5 and 7) presents cues for the next frame in the motion sequence at all times. **Autonomous** guidance is not shown as it operates independently of the trainee's actions.

ABSTRACT

Extended reality (XR) technologies are highly suited in assisting individuals in learning motor skills and movements—referred to as motion guidance. In motion guidance, the “feedforward” provides instructional cues of the motions that are to be performed, whereas

the “feedback” provides cues which help correct mistakes and minimize errors. Designing synergistic feedforward and feedback is vital to providing an effective learning experience, but this interplay between the two has not yet been adequately explored. Based on a survey of the literature, we propose design space for both motion feedforward and corrective feedback in XR, and describe the interaction effects between them. We identify common design approaches of XR-based motion guidance found in our literature corpus, and discuss them through the lens of our design dimensions. We then

discuss additional contextual factors and considerations that influence this design, together with future research opportunities for motion guidance in XR.

CCS CONCEPTS

- Human-centered computing → Visualization theory, concepts and paradigms; Mixed / augmented reality; Virtual reality.

KEYWORDS

Design Space, Extended Reality, Motion Guidance, Visualization

ACM Reference Format:

Xingyao Yu, Benjamin Lee, and Michael Sedlmair. 2024. Design Space of Visual Feedforward And Corrective Feedback in XR-Based Motion Guidance Systems. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA*. ACM, New York, NY, USA, 15 pages. <https://doi.org/10.1145/3613904.3642143>

1 INTRODUCTION

Augmented Reality (AR) and Virtual Reality (VR), collectively known as Extended Reality (XR), have gained significant attention in recent years for their potential to enhance training and performance in various domains. One particularly promising application of XR is in the domain of motion guidance systems, which assist individuals in learning and achieving specific motor skills, postures, or movements. These systems have the potential to make a positive impact in a wide range of fields, from sports [9] and healthcare [48] to industrial applications [42]. By presenting guidance and feedback around or even on the trainee's body, these systems avoid sloppy or incorrect performance of motions in the absence of instructor supervision [59]. The complex nature of XR technologies has opened up a vast design space for creating novel and effective guidance systems, as compared to previous technological setups involving flat 2D displays and pointing devices [3]. This design space is crucial for addressing key questions that arise when designing XR-based motion guidance systems, such as how the motion guidance is presented, how the system responds to user interaction, and how to help users rectify errors.

In this paper, we explore the visual design space of XR-based motion guidance systems. A complete guidance system simulates an environment for the trainee to practice or learn physical motions through XR. For this purpose, such a system should encompass two key components: motion feedforward, providing motion instructions before the trainee has taken action; and corrective feedback, aiding in behavior correction and performance improvement in response to the trainee's actions. Hence, we structure our investigation following the same two components. Recently, Diller et al. [13] conducted a related survey to summarize the work of using XR to provide corrective feedback for motor training. However, their survey focuses only on corrective feedback, and does not consider the relationship between it and the presented motion feedforward. Thus, building upon their survey, we consider design dimensions for both motion feedforward and corrective feedback, and explore the potential interplay between them.

As the foundational basis of our work, we curate and review a corpus of 38 papers on XR-based motion guidance systems based

on existing literature surveys and following citation trails (Section 3), forming the foundational basis of our work. For motion feedforward, we discuss design choices in the level of indirection, interactive update strategies, viewing perspectives, and any additional context cues (Section 4). For corrective feedback, we discuss design choices in information level, temporality, placement, and presentation methods (Section 5). Furthermore, we review technical setups of existing XR-based motion guidance systems from the literature to provide insights into the current state of the field with respect to our proposed design space (Section 6). We also illustrate how to utilize our design space to generate the appropriate XR systems based on two different scenarios (Section 7). We then discuss the broader context, constraints, and factors that influence the design of XR-based motion guidance systems, intending to show how to use our design space under different practical conditions (Section 8).

The goal of this work is to help create effective and user-friendly XR-based motion guidance systems that can support training, rehabilitation, and skill acquisition across various disciplines. As such, the contributions of this paper are as follows:

- Design space for motion feedforward and corrective feedback in the context of XR-based motion guidance.
- Analysis of common configurations and designs of motion feedforward and feedback based on a corpus of 38 papers, and identification of multiple underexplored research opportunities in the literature.
- Illustrative examples of using the proposed design space to create new XR-based motion guidance systems.
- Discussion of the related constraints and factors that further influence the design of motion feedforward and feedback.

2 BACKGROUND & RELATED WORK

2.1 XR-Based Motion Guidance Systems

Motion guidance systems provide instructions to assist trainees in performing specific motions in situations where on-site guidance from a human instructor is unavailable. This field covers a wide range of topics, including sports training [60], physical rehabilitation [52], and dance [21]. In this paper, we want to focus on motor skill training and dissociate this work from complex tool handling [6], assembly tasks [42], or delicate professional skill training [41]. Elsayed et al. [15] classified motion guidance into three main groups: (1) **posture guidance**, which involves providing clear visualizations of target postures at specific frames; (2) **path guidance**, which focuses on directing trainees' body movements along predefined paths in a 3D space; (3) **movement guidance**, which offers clear posture visualizations at each progressive frame along the path. In this paper, we will comprehensively explore motion guidance that encompasses all of the aforementioned concepts in the context of XR, and reclassify XR-based motion guidance based on our design dimensions (Section 4.5).

To achieve effective training, a motion guidance system needs to adequately inform the trainee what motions to perform and seek to reduce the number of errors made. Thus, it needs to include both *feedforward* (Section 2.2) and *feedback* (Section 2.3) mechanisms.

2.2 Feedforward

Feedforward, in a general context, provides users with information about available actions and expected outcomes before an interaction, guiding them on what to do and how to proceed. Muresan et al. [46] describe the three stages of how feedforward operates in a VR environment. First, the user needs to trigger a preview of desired actions. Subsequently, feedforward presents a preview to the user, showing the actions and their outcomes that can be taken at specific locations for specific objects. Then, the user can exit the preview and proceed to perform the action.

In the field of motion guidance, Sodhi et al. [51] defined feedforward as components that provide users with information about the shape of a motion before it is executed. This information is oftentimes required for trainees to properly learn and complete specific gesture sequences [3, 12, 18]. Feedforward visualization in motion guidance can take many forms, including motion paths [12], arrows [51], user skeletons [1], hand shadows [18], and body outlines [22]. Feedforward may also take the form of playful and funny cues to increase engagement, also known as “serious games”, which are often employed in the field of physiotherapy and rehabilitation [4, 10]. However, in this paper we focus primarily on task-based scenarios that provide clear movement instructions, focusing on the execution of simple, repetitive exercises rather than gamely motivational aspects [5].

2.3 Feedback

While feedforward provides guidance on the necessary steps to attain a specific result, *feedback* assists the user in comprehending what has happened or is currently happening. Feedback mechanisms specifically present users with the outcomes of particular interactions that they have made, thus relying significantly on users’ past actions and movements within the system [3].

In the field of motion guidance, there may be two types of feedback: corrective feedback and assessment feedback. Corrective feedback is designed to identify and rectify trainee errors or inaccuracies, thereby facilitating the improvement of knowledge and skills. For motion guidance, corrective feedback is employed when the deviation between the primary object (e.g., trainee’s joints [53] or a surgical saw [41]) and its target position exceeds a “threshold” pre-defined by the trainer or designer [31]. Corrective feedback can be presented visually [26], tactiley [30], auditorily [8], and even multimodally [49]. While this presents an interesting array of design choices, we focus primarily on visual feedback as it is the baseline mode that almost all motion guidance systems rely on—especially in XR. Regardless of the presentation mode, corrective feedback should be at least capable of indicating when or where errors are occurring or have occurred. In contrast, assessment feedback provides an overall “score” to the trainee after each trial is conducted, providing a quantifiable measure of their performance. We discuss this in more detail in Section 8.2.

2.4 Design Space of Motion Guidance

Several works have presented design space of how feedforward and feedback can be presented to users. An early work by Bau and Mackay [3] proposed a design space for classifying feedforward and feedback mechanisms in gesture training systems. They derived two

design dimensions regarding feedforward (level of detail and update rate) and four for feedback (recognition value, filtering, update rate, and representation). Despite using mouse and pen input and not XR, their work is still of high relevance to any motion guidance reliant on vision. Building upon this, Delamare et al. [11] identified four common design dimensions applicable to both feedforward and feedback mechanisms within gesture guidance systems, encompassing temporal, content, medium, and spatial characteristics. Their work explores the breadth of gesture guidance designs, including gesture recognition and trigger mechanisms for feedforward and feedback, but does not focus on the visual channel—especially in XR. Muresan et al. [46] explored the design of feedforward in VR based on its three key stages (Section 2.2), but focused primarily on physics-based interactions with everyday objects, which involve longer and more complex action sequences compared to motor skill training.

Most relevant to our work is the survey by Diller et al. [13] of visual corrective feedback cues in XR regarding motion training. They categorized visual corrective feedback in existing motion training designs based on nine attributes: MR technologies, point of view, abstraction type, temporal order, stages of learning, publication venue, body parts, use case, and visual cues. Their classification of visual cues, while suitable for describing common approaches in the literature, does not adequately describe and differentiate between the fundamental properties of each cue. For instance, “rubber bands” and “arrows” are very similar in that they both indicate a direction, with the main difference being the presence of an arrowhead. In contrast, we believe that the induction of visual cues will contribute to a better understanding of the nature of corrective feedback, and help further explore the relationships between visual cues and other elements within the system. Moreover, their descriptions of corrective feedback also include how future movements are presented (e.g., “guidance” in “abstraction type” and “upcoming” in “temporal order”), which should instead be classified separately as feedforward. We thereby extend their work to include the exploration of feedforward and corrective feedback in XR separately, the clarification of their functional divisions and connections, and further the analysis of the interplay between these two mechanisms.

In summary, previous work has yet to provide a holistic overview of visual design for XR-based motion guidance. They are either aimed at non-motion guidance scenes, or they do not use XR to present feedforward, and hence lack comprehensive consideration of both feedforward and feedback in the XR context. Our work aims to fill this gap through a literature search combined with our experience in related design. Concretely, we focus on motor skill training with simple and repetitive motions in task-based scenarios, where feedforward and feedback are both presented through XR and mainly on the visual channel.

3 METHODOLOGY OF LITERATURE REVIEW

To ensure that our design space is properly grounded in prior XR-based motion guidance research, we decided to conduct a literature review. It is through this process that we could brainstorm, define, and validate our proposed design dimensions for both feedforward and feedback.

To construct our corpus, we employed the snowballing methodology described by Wohlin [57]. The start set of our snowball comprises three logical subsets. *Start set A* contains two comprehensive survey papers: one on corrective feedback cues by Diller et al. [13], and one on rehabilitation exercises in XR by Butz et al. [5]. *Start set B* contains prominent publications (over 100 citations on Google Scholar) that fit our scope: Just Follow Me [58], ShadowGuides [18], LightGuide [51], YouMove [1] and Physio@Home [53]. *Start set C* contains publications within the past 5 years on XR-based motion guidance across a broad research spectrum, including: Kodama et al. [32] on training using virtual co-embodiment, Zhou et al. [61] on motion guidance with an MR mirror, Lilija et al. [39] on correction on virtual hand avatar movements, Yu et al. [59] on the influence of perspective in motion guidance, and Dürr et al. [14] on the virtual appearance of feedforward.

To expand the corpus, we employed the backward snowballing method on *start set A*, while concurrently applying both forward and backward snowballing methods on *start set B* and *start set C*. For each paper to examine, we first conduct a preliminary screening based on its title, abstract and keywords, and then decide on inclusion or exclusion by reading its system design and comparing it with our intended scope. After including a paper in our corpus, we applied both forward and backward snowballing, and repeated this process until no new papers in its associated citation chain were left. We only included papers published after 2002, which was the year when the Just Follow Me system [58], the first system to employ VR for motion training, was published.

We opted against using a database query approach for two main reasons. First, there exist many synonyms for motion guidance, thus making it challenging to search by. For example: motion, motor, movement, and exercise; and guidance, tutorial, training, and assistance. Second, a database query may inadvertently exclude XR-based motion guidance in specific application domains such as medical and sports science, especially when limiting the search to certain publishers and/or venues (e.g. ACM, IEEE).

Our final corpus consists of 38 papers, which includes the papers in our start set. Among them, 16 were published in venues related to XR and visualization (e.g., IEEE VR, ISMAR, TVCG; ACM VRST; Springer VR), 15 in human-computer interaction (e.g., ACM CHI, UIST, IUI, ICMI, ITS, AH; NordiCHI), and the remaining 7 in other domains such as gaming (IEEE Trans. Games), artificial intelligence (Frontiers in Robotics and AI; ACM TIST), and communication technology (Frontiers in ICT). Table 2 shows all of the papers in our corpus. While this corpus might not be exhaustive, we believe that it is nonetheless a representative sample of the space.

We use this corpus to devise design space for motion guidance in terms of its two components: motion feedforward (Section 4) and corrective feedback (Section 5). We first extracted initial design dimensions from existing literature [3, 11, 13]. We then augmented these dimensions by examining the many different technical setups present in our corpus, drawing on our own experience in the fields of XR and motion guidance. We iteratively refined the design dimensions through many discussions among the authors, aiming to avoid omissions, ambiguity, and overlap between dimensions. Validation was performed at each stage by applying the design dimensions to all papers in our corpus until a final design space was agreed on.

4 DESIGN SPACE FOR MOTION FEEDFORWARD

In a motion guidance system, the feedforward provides information on how to move or perform specific motions. It is the instructions, cues, or guidance given to trainees *prior* to the course of a movement, which can help them understand the desired motion and how to execute it properly, such as the motion path [3] or the visualization of postures [18]. We further refine the design space of feedforward in XR-based motion guidance based on the two original dimensions defined by Bau and Mackay [3]: level of detail (the granularity of the motion sequence presented in feedforward), and update rate (the frequency that feedforward updates). We adapt these two dimensions to XR as follows. First, the immersive and embodied nature of XR allows motion information to be represented in a very direct and life-like form as compared to on a desktop in Bau and Mackay's [3] work, which we capture as the **level of indirection** of the feedforward. Second, the highly interactive nature of XR has us consider how the feedforward might respond to the trainee's actions throughout the motion, rather than being an independently updating entity, thus resulting in our **interactive update strategy** dimension. We also consider the broader scope of possibilities to visually present information in XR, which includes both the **viewing perspective** between the trainee and the feedforward, and any **additional contextual cues** that are associated with the motion. We further discuss the interactions and possible configurations between these dimensions.

4.1 Level of Indirection

XR is capable of bringing information into forms that are either inherently abstract, or forms that bear a more direct resemblance to reality [36]. The same is true for motion guidance, as the feedforward is capable of either showing exactly what motion should be performed, or instead an (intentionally) vague representation of it. We refer to this as the **level of indirection** of the feedforward's visual representation. We describe three levels, but note that this should be treated more as a spectrum with no strict boundaries between them (Figure 2a).

Explicit. The feedforward uses a low level of indirection such that it explicitly represents the desired motion with little to no ambiguity. Usually represented as a 3D graphic, it enables the trainee to simply observe the feedforward and imitate its motion or posture, thus providing no room for misinterpretation. For instance, AR-Arm [22] uses transparent 3D arms which provide direct instruction for the trainee to simply overlap their arms with the feedforward. Explicit feedforward need not be realistic in appearance however. For instance, YouMove [1] uses stick figures as its feedforward to show full-body postures. It is still obvious to the trainee that they should move and align their own skeletal posture with the feedforward. Therefore, a low cognitive cost is associated with explicit guidance, as trainees can simply follow what they see.

Implicit. The feedforward uses a medium level of indirection such that it still shows the desired motion, but requires some amount of interpretation and inference by the trainee to accurately perform. A classic example of this is motion guidance systems which use trajectories to show movement paths that need to be

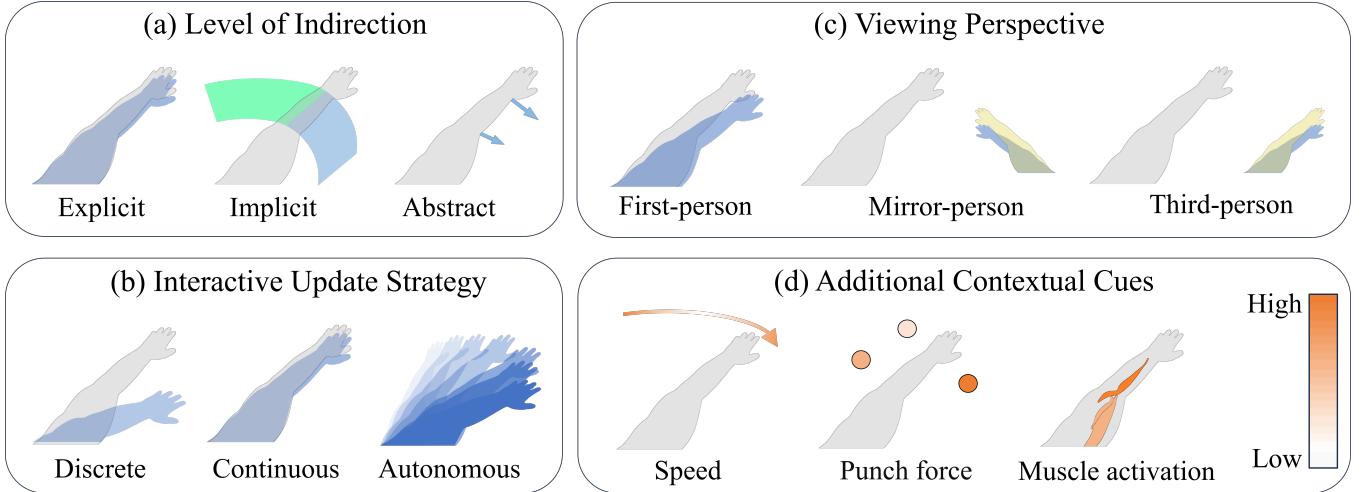


Figure 2: Design dimensions of motion feedforward. (a) level of indirection at *explicit*, *implicit* and *abstract*. (b) interactive update strategy at *discrete*, *continuous* and *autonomous*. (c) viewing perspective at *first-*, *mirror-* and *third-person* perspectives; in the latter two perspectives, the feedforward will show around the duplicate of the trainee’s avatar. (d) additional contextual cues such as *speed*, *punch force* and *muscle activation*. For illustrative purposes, grey represents the trainee’s avatar in an egocentric perspective, blue represents feedforward instructions, green represents the completed part in *implicit* guidance, yellow represents the dynamic duplicate of trainee’s avatar in *mirror-* or *third-person* perspective, and orange represents severity of speed, punch force, or muscle activation.

followed [52, 53, 59]. While the trajectory, when seen in its entirety, does not present some motion or posture to imitate, it is the traversal from one end of the trajectory to another that results in a guided motion. The need to interpret the feedforward may also be beneficial, as it provides the trainee with more freedom in how they perform the motion. For instance, Kosmalla et al. [33] use a projector to highlight hand holds while rock climbing. While this indicates which holds to move their limbs to, it does not indicate **which** limb to move, thus allowing for more self-decision-making. This decision-making naturally incurs a higher cognitive cost as trainees need to think for themselves the best way to make the desired motion.

Abstract. The feedforward uses a high level of indirection such that it loosely describes the desired motion, but does little to show how it exactly should be performed. This results in a high level of ambiguity in which the trainee needs to approximate the motion as best they can. An example of this is LightGuide [51] which projects visual cues onto the trainee’s hand such as directional arrows. These arrows, in particular, hint at the trainee a general direction in which their hands should be moved, but with little indication of how far, how fast, or even how straight to move. Alternatively, the trainee may be required to perform the motion purely by memory, with the feedforward serving only as a reminder of the motion to perform. Such use would primarily be to test the trainee’s knowledge of already practiced motions, with the lack of guidance being intentional [32, 39]. This clearly has the highest cognitive cost due to trainees not being able to rely completely on the feedforward.

4.2 Interactive Update Strategy

For complex 3D motions (e.g., long, intersecting, or overlapping movements), visualizing all postures of the entire motion sequence may be too distracting due to visual clutter [59]. Hence, motion guidance systems often employ checkpoints throughout the motion sequence, progressively updating or revealing parts of the feedforward in response to the trainee’s movement as each checkpoint is reached—when their body aligns with the checkpoint’s target posture. Systems can vary in their **interactive update strategy** (i.e., the distribution of the checkpoints), categorized as either *discrete* guidance, *continuous* guidance, or *autonomous* guidance, as illustrated in Figure 2b.

Discrete Guidance. Discrete guidance sets checkpoints only for frames related to critical postures. The trainer determines the critical postures in the motion sequence, which trainees should perform as accurately as possible. When the trainee performs critical posture, the checkpoint is met and the next stage of feedforward is shown (and the previous one hidden). This approach can add greater emphasis on crucial moments of the motion, allowing trainees to concentrate on accurately performing the critical postures [26]. The effectiveness of discrete guidance is naturally influenced by the choice of critical postures. Alternatively, discrete checkpoints can subdivide a motion into more semantically meaningful components, which may assist trainees in remembering this sequence. For instance, consider a complex ballroom dance sequence with many dance steps performed in rapid sequence [21].

Continuous Guidance. Checkpoints are distributed throughout every single frame of the motion sequence [22], thus providing

a constantly updating flow of visual cues that guide the trainee throughout the entire motion. This is a highly responsive form of motion guidance, in which the feedback continuously adapts to the trainee's progress and posture alignment, allowing for real-time adjustments and skill refinement. However, trainees need to visually focus on the constantly updating feedforward, which may lead to a higher dependency on this guidance [39].

Autonomous Guidance. In comparison to the other two strategies, autonomous guidance sees no checkpoints being used, and instead has the feedforward automatically update on its own regardless of the trainee's input. Such guidance is typical of videos or 3D animations [14, 55], where trainees can observe the complete motion sequence before attempting it themselves. However, the lack of responsiveness of the feedforward to the trainee's own movements may result in a less engaged learning experience. This may also make it harder to fully grasp the intricacies and nuances of the motion without further context or explanations (e.g., assistive textual cues). However, autonomous guidance is well-suited for presenting motions that need to be performed at a specific speed, be it fast or slow, as the feedforward can show this in real-time.

It is possible to combine multiple interactive update strategies together. For instance, the *GT_PAUSING* technique by Dürre et al. [14] uses both *discrete* and *autonomous* guidance by playing short animations between critical postures of arm motions.

4.3 Viewing Perspective

XR research often considers how a user's perspective of the (virtual) world and its content influences their behavior, and this is no different for motion guidance [15, 59]. By varying the **viewing perspective** of the trainee relative to the feedforward, their perception of both the feedforward and of their own body changes, thus influencing their understanding and execution of the desired motions (Figure 2c). We describe three common viewing perspectives in the literature: *first-person*, *mirror*, and *third-person*.

First-Person (1PP). This perspective tailors the feedforward to be viewed from the viewpoint of the trainee, immersing them directly in the context of the motion, and thereby providing a more natural understanding of it. A common design is to use a transparent ghost arm that the trainee should follow with their own arms [14, 22]. This encourages a sense of embodiment and kinesthetic empathy, enabling trainees to closely align their actions with the guidance cues. However, motions involving multiple limbs at opposing parts of the trainee's body are not well suited for first-person perspectives, as they would need to frequently turn their heads to see the required feedforwards, thus increasing fatigue [59].

Mirror (MPP). Feedforward can be overlaid on a reflection of the trainee's own body viewed on a (virtual) mirror. Depending on the technological setup, the appearance of such reflections may be consistent with their real bodies (screen-based [53] or MR mirror [61]), or an avatar [55]. The mirror perspective naturally allows the trainee to see their entire body and the feedforward all at once. It can also simulate the trainee's surroundings to provide a meaningful context for a better training experience, such as a wall-size mirror in a dance studio for ballet training [1].

Third-Person (3PP). This perspective situates the trainees as an external observer of the feedforward, analogous to watching someone else perform the motion. In XR, this is typically a virtual coach using pre-recorded motion sequences such as to learn Tai-Chi [23], but can also be the virtual avatar of a remote trainer providing live motion guidance [26]. While distancing the trainee from the motion itself, the third-person perspective provides an exocentric view of the entire motion, thus promoting a more complete understanding of all facets of the motion—especially for motions involving multiple limbs [15].

4.4 Additional Contextual Cues

All XR-based motion guidance systems, by definition, should provide feedforward instruction of the motion sequence to perform. In some contexts however, further information relevant to the motion may need to be provided to the trainee. One such information is that of speed. While the feedforward itself, particularly in autonomous guidance, can show the speed outright, **additional contextual cues** can provide this information instead (Figure 2d). For example, Yu et al. [59] used a small sphere that moves along the motion path as an indication of speed, whereas Gutierrez et al. [21] encoded the speed of the current ballroom dance segment through the size of solid arrows on the ground. Another form of contextual information is that of muscle activation. This can be seen in the work by Zhu et al. [62], which overlays colors on a virtual coach's muscles to indicate to the trainee which muscles are being activated in leg rehabilitation exercises. Of course, it is up to the trainer to decide what additional information they present to the trainee and what cues they use, as this may range from the aforementioned speed and muscle activation, to even the facial expressions to make or the lyrics to be sung at different parts of a dance routine.

4.5 Configurations & Commonalities

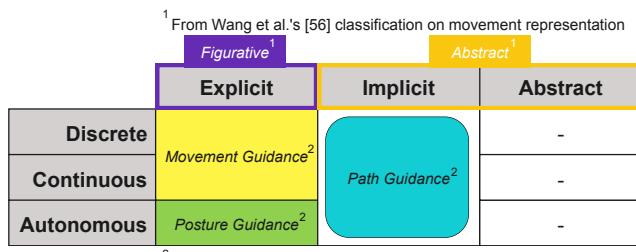
Among the four dimensions described in this section, **level of indirection** defines how to visualize motion cues at a certain moment, while **interactive update strategy** determines how this visualization will change with the progression of the training. Therefore, these two dimensions collectively present the fundamental information of the motion sequence, giving rise to nine potential configurations of motion feedforward. These configurations are capable of describing all of the feedforward designs in our literature corpus. Descriptions of each configuration and the papers that use them are presented in Table 1. In addition, Figure 1 provides an illustrative view of how *continuous* and *discrete* guidance can respond to trainee interaction (*autonomous* is excluded due to its independent operation from trainee interaction). To further verify the orthogonality of the two dimensions, we developed basic prototypes of each of the 9 configurations to demonstrate how they may function. Videos showcasing these prototypes are included in the supplementary material.

We further validate our nine configurations by comparing them with two existing classifications of motion guidance and representations, and show how they overlap with each other. This is shown graphically in Figure 3. First, the classification of rehabilitation movement representation by Wang et al. [56]. Their “abstract”

Table 1: Possible configurations of motion feedforward and the papers which use them. Illustrations are in Figure 1.

Configuration	Explicit	Implicit	Abstract
Discrete	DisExp: Explicit representation of desired motions, which updates at critical frames of the motion sequence. Commonly used for posture guidance, with less emphasis on the movements between each critical posture [1, 2, 7, 9, 14, 15, 18, 21, 26, 28, 34, 40, 47–49, 54, 59].	DisImp: Implicit representation of desired motions requiring interpretation, which updates at critical frames of the motion sequence. Used for when the trainee can perform the critical posture in any way they wish [33].	DisAbs: Approximate representation of desired motions, which updates at critical frames of the motion sequence. Commonly used to indicate the starting point of a motion when testing short-term retention [32, 39].
Continuous	ConExp: Explicit representation of desired motions, which updates throughout the motion when trainees satisfy the postural accuracy. Commonly used for slow, high accuracy motions [22, 29, 31, 39].	ConImp: Implicit representation of desired motions requiring interpretation, which updates throughout the motion when trainees satisfy the postural accuracy. Commonly used with motion trajectories, where concrete postures are interpreted as the boundaries between completed and uncompleted parts [17, 20, 52, 53, 59].	ConAbs: Approximate representation of desired motions, which updates throughout the motion when trainee satisfies the positional accuracy [51, 54].
Autonomous	AutoExp: Explicit representation of desired motions, which is autonomously played without interaction with the trainees. Commonly used as animations or videos demonstrating the desired motion(s) [1, 14, 16, 19, 23, 27, 32–35, 44, 50, 55, 58, 60].	AutoImp: Implicit representation of desired motions requiring interpretation, which is autonomously played without interaction with the trainees. A possible example is a motion trajectory that autonomously progresses the boundary between completed and uncompleted parts to indicate the desired speed of motion.	AutoAbs: Approximate representation of desired motions, which is autonomously played without interaction with the trainees. A possible example is a sequence of abstract feedforward cues that indicate the order of motions to perform (similar to a <i>Simon Says</i> game).

representation contains motion paths and directional cue representations such as arrows, which correspond to our *implicit* and *abstract* levels of indirection respectively. Their “figurative” representation refers to movements presented as a humanoid body, which should be covered by our *explicit* feedforward. Second, the classification by Elsayed et al. [15] on different types of motion guidance, already described in Section 2.1. In it, “posture guidance” refers to *discrete explicit*, “path guidance” refers to a subset of *implicit* guidance, and “movement guidance” refers to a combination of *continuous explicit* and *autonomous explicit* guidance.

¹ From Wang et al.'s [56] classification on movement representation² From Elsayed et al.'s [15] classification on motion guidance**Figure 3: Diagrammatic overview of how our design dimensions of level of indirection and interactive update strategy for feedforward relate to the original classifications by Wang et al. [56] and Elsayed et al. [15].**

5 DESIGN SPACE FOR CORRECTIVE FEEDBACK

In a motion guidance system, the corrective feedback provides alerts, reminders, and advice to trainees about the errors they make during training. To better investigate this space, we classify corrective feedback into two major groups: INTUITIVE and ADDITIONAL.

INTUITIVE. INTUITIVE feedback is borne out of the design of the feedforward. Feedforwards that are *explicit* (and sometimes *implicit*) provide a low enough **level of indirection** such that trainees can immediately notice disparities between their movements and the feedforward. This binary perception of errors therefore serves as an instinctual indication that the motion is not performed correctly, even in lieu of a quantification of this deviation. Hence, this phenomenon has been regarded as a form of corrective feedback within a number of related literature. In the survey by Diller et al. [13], the visual cues classified as “end position”, “transparent target avatar”, “opaque target avatar”, “video overlay”, “body outline”, “movement abstraction” and “trajectory” all fall into this group, with the first five cues manifesting themselves through postural differences, while the last two reflecting trajectory differences. Likewise, systems that use *continuous* [22] or *discrete* [18] guidance naturally indicate errors due to the need to align the trainee’s body to the required checkpoint. If the alignment criterion is not met, the feedforward will not visibly progress, thus prompting the trainee

to realize that an error has occurred and that they should correct it. Conceptually, these two facets harmonize with each other, forming a symbiotic relationship.

ADDITIONAL. ADDITIONAL feedback can also be employed to overcome some of the limitations that INTUITIVE feedback has. For example, trainees may not even notice that their body is misaligned with the feedforward, particularly in multi-limb motions or errors made in the depth direction [15]. Therefore, considering ADDITIONAL corrective feedback (visual cues like color and arrows) can still be beneficial, which requires additional visual encodings.

We build our design dimensions based on the survey by Diller et al. [13]. First, to remove the ambiguity about feedforward from our design space of feedback, as mentioned in Section 2.4, we exclude the feedforward elements from their attributes, namely “upcoming” from “temporal order” and “guidance” from “abstraction type”. Then, we take “temporal order” as our **temporality (when)** dimension, and we include their “abstraction type” and “visual cues” as our basis to explore the **presentation (how)** dimension. Finally, we propose dimensions of **information level (what)** and **placement (where)** to complete the description of corrective feedback visualization. We believe that this structured division avoids overlap among dimensions, and that assigning descriptiveness to dimensions facilitates a more detailed analysis of feedback and the relationship between feedback and feedforward.

In contrast, since INTUITIVE feedback is presented through the visual differences between the trainee’s movement and feedforward rather than visual encoding, its design choices are limited in the **what**, **when**, **where** dimensions. Therefore, we summarize its visual designs only in the **how** dimension.

5.1 What: Information Level

The first aspect of corrective feedback is **what** information it tells the trainee, which naturally influences the ease and extent to which they can make corrections. In motion guidance, corrective feedback typically falls under three progressive levels of information: *detection*, *magnitude*, and *rectification*.

Detection. As the foundational level of corrective feedback, detection provides a straightforward binary output, conveying whether the trainee’s body conforms to the required motion or not. Existing works show various approaches in how to encode this output, such as text to display the alignment status of each of the trainee’s skeletal joints [7], or color on the feedforward itself switching between red and green [26]. This information level requires the trainer to determine the ideal threshold which dictates the binary switch. Too low a threshold increases the demanded accuracy but is prone to “flickering”, too high a threshold decreases the demanded accuracy but may make for sloppier trainee movements. Moreover, detection does not provide any further detail about the error or targeted improvements to be made.

Magnitude. The second level of corrective feedback, quantifies the magnitude and extent of the deviation from the target posture, referred to as deviation distance. This distance is often mapped to visual channels that allow trainees to pre-attentively notice subtle mistakes in their moment, such as transparency [29]. Magnitudes can also be seen used in post-hoc analysis after the training process,

with the extent of errors being marked on the motion path [50, 52]. This information level requires the designer to determine the output visual ranges to map the deviation distance to. That is, at what extents does the given deviation distance result in the lowest and highest intensity of visual output.

Rectification. As the third and highest level of corrective feedback, rectification provides trainees with direct guidance and instructional cues to correct and refine their motions. In other words, it tells the trainee “how” to correct their mistakes. For instance, Yu et al. [59] used lines connecting the trainee’s arm to the corresponding target positions on the feedforward. These lines serve a dual purpose: their length indicates the deviation distance, and their direction cues the trainees as to how to adjust their arm to correct the deviation. Alternatively, rectification information can be conveyed by the textual suggestions after the training process. For instance, VCoach [9] gives comments to trainees about their performance and suggestions regarding punch direction and speed.

5.2 When: Temporality

Corrective feedback can differ in terms of **when** it is actually presented to trainees. This includes two main choices: *real-time* and *post-hoc* feedback.

Real-Time. This is the instantaneous delivery of feedback to trainees as they are engaged in the motion training. This allows trainees to promptly identify any errors in their performance and make timely corrections, thereby creating a rapid learning feedback loop. Examples include showing mismatches in postures [55] and the immediate highlight of movement errors [20].

Post-Hoc. This delays the delivery of feedback until only after trainees have completed a motion trial. While post-hoc feedback lacks the immediacy inherent to real-time feedback, it provides a longer window for self-reflection, allowing trainees to thoroughly analyze their errors and develop effective strategies for improvement. For instance, post-hoc analysis described in [16] provides trainees with a review of the movement accuracy over time in the form of a graph, aiming to improve the trainee’s movement accuracy at corresponding time points.

5.3 Where: Placement

Where the corrective feedback is placed directly influences not only the trainee’s ability to make corrections, but also their ability to see the feedforward itself. A misplaced or distracting corrective feedback may unintentionally harm and not help the trainee’s performance. We outline three strategies of where corrective feedback can be placed that are common in prior research.

Motion Feedforward. The feedback is embedded directly into the feedforward itself, reducing the need to shift attention between the two. This augments the contextual relevance, enabling trainees to promptly identify errors and refine their motor skills all while visually focusing on the feedforward to execute the motion. An example can be found in the OneBody system [26], where the skeletal joints on the feedforward change color in response to the misalignment of the trainee’s body to it (Figure 4a).

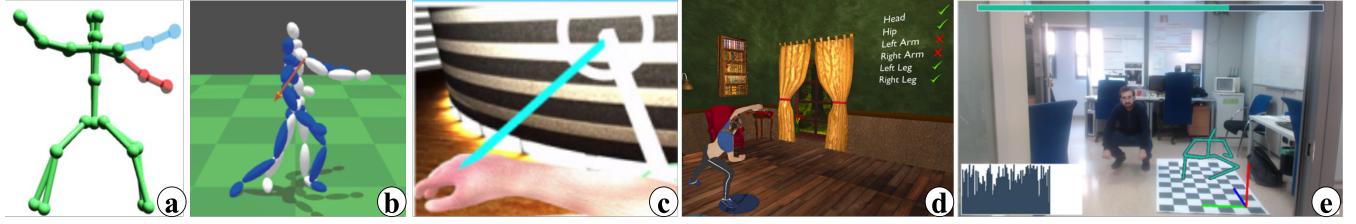


Figure 4: Examples of corrective feedback in the literature: (a) *Color*: both the trainee’s avatar and the virtual trainer are represented as ball-and-stick models, changing to green when corresponding body parts align, or to red and cyan when they do not align [26] (courtesy of Vetere, ©ACM). (b) *Direction*: a red arrow is positioned on the trainee’s shoulder (in white), and provides a *rectification* feedback by pointing at the virtual coach’s (blue) shoulder [47] (courtesy of Oshita, ©IEEE). (c) *Size*: the length of cyan lines indicates the deviation distance between the trainee’s wrist and the desired posture [59] (courtesy of Yu, ©IEEE). (d) *Text*: the *detection* of the alignment status of each body joint attached to a wall [7] (courtesy of Caserman, ©IEEE). (e) *Graph*: the graph on the bottom-left corner presents the *magnitude* of the movement error throughout the entire training phase [16] (courtesy of Escalona, ©Springer).

Trainee’s Body. The feedback is anchored to the trainee’s own body, fostering a heightened connection between the feedback and the trainee’s self-embodiment. Here, the “body” may be the trainee’s own physical body when using AR, virtual avatar while in VR, or the dynamic duplicate in the *mirror-* or *third-person* perspective. For instance, Oshita et al. [47] show how, during post-hoc analysis, arrows embedded into the trainee’s avatar can be used to point towards the correct position corresponding to the feedforward. The advantage of this approach is that trainees might perceive the feedback as an extension of their own actions, leading to quicker adjustments in their movements. However, this strategy may introduce visual clutter and distract away from the feedforward [59], thus requiring careful design in balancing their saliency while not being overly intrusive (Figure 4b).

Environment. The feedback is projected onto the surrounding environment away from the trainee and the feedforward, thus minimizing undue attentional distractions while performing the motion. Moreover, this frees up the space in the trainee’s vicinity for the feedforward to convey a greater amount of information. As an example, Caserman et al. [7] position corrective feedback as text logs on the walls of the virtual environment (Figure 4d). However, the challenge lies in ensuring the feedback is still clearly visible, particularly for real-time feedback, as it introduces another visual element that the trainee needs to pay attention to.

5.4 How: Presentation

Lastly, **how** the corrective feedback is presented to trainees is important. This refers to the methodological visual representation of movement errors and instructional cues. We draw upon the 13 visual cues summarized in Diller et al.’s survey [13] as a reference, but we also streamlined and consolidated them to explore common visual features and corresponding principles within these cues.

To convey error information and corrective suggestions in a timely manner, corrective feedback should be designed to capture sufficient attention among other visual elements (feedforward, avatar, and surroundings). Therefore, we employ the pre-attentive mechanism [24] (also known as the visual pop-out theory [45]) to

analyze visual cues. Specifically, we abstract the cues in [13] by the visual features based on their pre-attentive mechanism.

As mentioned at the start of Section 5, INTUITIVE feedback is borne from the trainee being able to perceive misalignments between the visual feedforward and their own movements at the same moment. Diller et al.’s [13] survey describes several cues that are, we argue, designs of motion feedforward that take advantage of this INTUITIVE feedback as a corrective mechanism. This includes the *shape* of movement for the “*trajectories*” cue, and the *positions* of body joints and the *orientations* of limbs for cues including “*transparent target avatar*”, “*opaque target avatar*”, “*video overlay*”, and “*body outline*”. The differences in these three features can be used to capture the trainee’s attention and correct for errors.

The ADDITIONAL feedback mechanism encodes visual attributes of an object to present feedback information. Such an object can be the feedforward representation itself [55] (which may serve in conjunction with INTUITIVE feedback), the trainee’s avatar [26], or other objects in the environment [16]. These visual attributes may take the form of *color*, *direction*, or *size*, and may also be reflected in the content of *text* or *graph*. They update as the trainee performs the motion, signifying to the trainee whether the movement is performed correctly.

Color. Color can not only serve as an intuitive indicator of movement accuracy (e.g., green denoting correctness and red denoting errors [48, 53]), but also enable the representation of the extent of deviations by interpolating color hues. For instance, a higher intensity of red signifies a higher *magnitude* of movement error [20].

Direction. Directional cues indicate where and how the trainee’s body parts should move to correct for alignment. Both the “*rubber bands*” and “*arrows*” in [13] employ this approach. The rubber band connects the actual and target body positions with a line segment [59], while the arrow, when an error occurs, appears at the trainee’s deviated body joint pointing towards the corresponding target position [47]. Another example can be found by projecting fan-shaped ripples onto the ground to correct the direction of footprint movement [50].

Size. Alterations in the length, area, or volume can effectively catch the trainees' attention [24, 25] and evoke a direct association with the *magnitude* of movement errors. For instance, the *length* of the connecting line between body joints of the trainee and feedforward directly visualizes the deviation distance [59]. Building upon this, Gebhardt et al. [20] replaced this line with a pair of opposing cones, encoding both their heights and base widths in terms of deviation distance between body joints, and eventually presented the movement error by the *volume* of the cones.

Text. Text feedback can serve to alert trainees to the occurrence of errors, such as the mismatch of body joints [7], or go further by providing direct corrective suggestions [47].

Graph. Graphs provide an abstract method of presenting movement information to the trainees and trainer (e.g., speed and deviation distance). It is well suited, for example, to provide a comprehensive visual overview of movement errors throughout a continuous motion sequence [16].

6 COMMON DESIGN APPROACHES OF MOTION GUIDANCE

With a corpus of 38 papers on XR-based motion guidance and a design space to describe them, we now delve into the exploration of its common design approaches. We were particularly interested in examining popular design choices and the commonalities between them. As such, Table 2 shows all 56 motion guidance setups found in our corpus, along with the distribution of their design choices and use cases in our design space. Note that as the design of **INTUITIVE** feedback is strictly based on the feedforward, we only list design choices related to **ADDITIONAL** feedback. To achieve a more comprehensive understanding of the literature, we categorized the use cases of all setups into four categories:

- ■ *Sports training*, in which the training process may involve high-intensity movements that may cause fatigue.
- ■ *Rehabilitation*, which aims at assisting in the recovery and healing of the body.
- ■ *Dance tutorials*, which aims to provide instruction on various dance techniques.
- ■ *Others*, which contains training that have no requirements for movement speed or have much higher requirements for movement accuracy than temporal performance.

After making numerous observations of how feedforward and feedback have been used together in XR-based motion guidance, we also identify potential open research questions and opportunities where relevant.

Autonomous feedforward is typically explicit. All 19 setups that employed the *autonomous* update strategy exclusively utilized *explicit* guidance. One possible reason is that autonomously played motion feedforward does not wait for the trainee to interact, leading to a shorter duration for the trainee to observe and comprehend individual frames of motion. Therefore, there is a preference for using *explicit* guidance to ensure efficiency in observation and imitation. Based on our description in Table 1, theoretically, *implicit* and *abstract* guidance can undergo *autonomous* updates, which is technically feasible yet remains unexplored in the literature.

Use case affects perspective choice. The frequency of **viewing perspective** varies across different categories of use cases. In *Sport Training*, *1PP* is the most prevalent, followed by *3PP*, with only one instance of *MPP*. Because these types of movements involve muscle strain or speed requirements, it is necessary to minimize coordinate system transformations related to motion direction—especially symmetric transformations—in order to reduce fatigue. While for *rehabilitation*, the use of the *1PP* is the least frequent (4 out of 13). Compared to *MPP* and *3PP*, *1PP* is generally more intuitive and beneficial for maintaining postural accuracy of 3D complex motions [59]. However, *rehabilitation* exercises are relatively simple and typically consist of 2D segments, which may not fully leverage the advantages of the *1PP*. Moreover, *MPP* and *3PP* are more commonly employed in traditional rehabilitation scenarios. Among the six cases of *dance tutorial*, five of them employed the *MPP* to instruct ballet in a simulated rehearsal studio, while the only ballroom dance tutorial projected cues of footwork on the ground in *1PP*. Lastly, within the 25 use cases in *others* category, *1PP* is the most frequently used, with a discontinuous style (16 out of 25), as these cases specifically emphasize the postural accuracy of 3D motions. Building upon these findings, we believe that despite the introduction of novel display and interaction techniques, the design of XR-based motion guidance should still consider as much of how trainees learn and practice motions in the traditional real-world context, especially in terms of perspective.

Intuitive feedback is induced by feedforward. Twenty-one setups do not use any **ADDITIONAL** corrective feedback. Of these, 19 used an *explicit* form of feedforward, suggesting that a low **level of indirection** is indeed sufficient to not require further forms of feedback. The two exceptions to this used *implicit* feedforwards. The work by Lilija et al. [39] focused on testing short-term retention of hand movements, and thus would not require corrective feedback in this context. The other on ClimbVis by Kosmalla et al. [33] is primarily motivated by presenting route of holds to follow. In addition, two special cases display corrective feedback associated with the feedforward in ways not described in Section 5.4. One case used a progress bar in the *environment* to indicate time spent maintaining a posture, which would reset when this posture is broken, thus indicating a *detection-level* corrective feedback [1, 2]. The second case occurs in LightGuide [51], where the *abstract* feedforward and *rectification* feedback are merged together onto an arrow projected onto the trainee's hand, which always points towards the endpoint of the motion regardless of hand position or errors made.

Non-rectification feedback requires explicit feedforward. Among the 18 cases where *magnitude* or *detection* feedback was employed, 15 employed *explicit* guidance, two used *implicit* guidance, and one provided *abstract* cue for testing short-term retention [32]. This was because the **ADDITIONAL** feedback in these cases did not explicitly instruct how to rectify movements, necessitating reliance on **INTUITIVE** feedback—the feedforward to correct the trainee's movements.

2D corrective feedback. A small number of setups utilized 2D forms of corrective feedback. *Text* was employed in three setups: in one case it was used as a form of *detection* to indicate unaligned joints [7], and in the two other cases it was used as *rectification* to

Table 2: 56 motion guidance setups assigned to each of our design dimensions from a corpus of 38 papers.

Feedforward			Feedback						Use Case	Reference	
Update	Indirection	Perspective	What Level	When Temporality	Where Placement	How Color	Direction	Text	Graph		
Dis	Exp	1PP	Rectification	RT	Environment			•			[9]
Dis	Exp	MPP									[28]
Con	Exp	3PP	Magnitude	PH	Trainee	•					[29]
Con	Exp	3PP	Magnitude	RT	Trainee	•					[29]
Auto	Exp	1PP									[33]
Dis	Imp	1PP									[33]
Auto	Exp	3PP									[35]
Dis	Exp	1PP	Detection	PH	Environment	•					Basketball Shot [40]
Dis	Exp	3PP	Rectification	PH	Trainee		•	•			Tennis [47]
Con	Abs	1PP	Rectification	RT	Trainee		•				Fitness [54]
Dis	Exp	1PP	Rectification	RT	Trainee		•				Fitness [54]
Auto	Exp	3PP	Detection	RT	Feedforward	•					Skiing [60]
Auto	Exp	MPP	Magnitude	PH	Environment				•		Physiotherapy [16]
Auto	Exp	MPP	Magnitude	RT	Environment				•		Physiotherapy [16]
Con	Exp	1PP									Taichi [22]
Auto	Exp	3PP									Taichi [23]
Dis	Exp	1PP	Detection	RT	Feedforward	•					Physiotherapy [26]
Dis	Exp	3PP	Detection	RT	Feedforward	•					Physiotherapy [26]
Con	Exp	MPP	Rectification	PH	Trainee		•				Rehabilitation [31]
Con	Exp	MPP	Rectification	RT	Trainee		•				Rehabilitation [31]
Auto	Exp	3PP									Rehabilitation [44]
Dis	Exp	MPP									Rehabilitation [48]
Auto	Exp	1PP	Rectification	RT	Environment		•		•		Gait Rehabilitation [50]
Con	Imp	1PP	Magnitude	RT	Environment	•					Rehabilitation [52]
Con	Imp	MPP + 3PP	Rectification	RT	Trainee	•	•		•		Physiotherapy [53]
Auto	Exp	MPP									Ballet [1]
Dis + Auto	Exp	MPP									Ballet [1]
Dis	Exp	MPP									Ballet [2]
Dis	Exp	1PP									Ballroom Dance [21]
Auto	Exp	MPP	Magnitude	PH	Environment				•		Ballet [34]
Dis	Exp	MPP	Magnitude	PH	Environment				•		Ballet [34]
Dis	Exp	MPP	Detection	PH	Environment	•		•			Yoga [7]
Auto	Exp	1PP									Musical Conductor [14]
Dis + Auto	Exp	1PP									Musical Conductor [14]
Dis	Exp	1PP	Detection	RT	Feedforward	•					Others [15]
Dis	Exp	3PP	Detection	RT	Feedforward	•					Others [15]
Con	Imp	1PP	Detection	RT	Feedforward	•					Hand Gesture [17]
Dis	Exp	1PP									Hand Gesture [18]
Auto	Exp	1PP									Writing [19]
Con	Imp	1PP	Rectification	RT	Trainee	•	•		•		Others [20]
Auto	Exp	3PP									Squats [27]
Auto	Exp	MPP									Squats [27]
Auto	Exp	1PP	Detection	RT	Feedforward	•					Others [32]
Dis	Abs	1PP	Detection	RT	Feedforward	•					Others [32]
Con	Exp	1PP									Others [39]
Dis	Abs	1PP									Others [39]
Dis	Exp	3PP	Rectification	RT	Trainee		•		•		Others [49]
Con	Abs	1PP	Rectification	RT	Trainee		•		•		Others [51]
Con	Abs	1PP	Rectification	RT	Trainee	•			•		Others [51]
Auto	Exp	MPP	Detection	RT	Feedforward	•					Others [55]
Auto	Exp	1PP									Writing [58]
Con	Imp	1PP	Rectification	RT	Trainee		•		•		Others [59]
Con	Imp	3PP	Rectification	RT	Trainee		•		•		Others [59]
Con	Imp	MPP	Rectification	RT	Trainee		•		•		Others [59]
Dis	Exp	1PP	Rectification	RT	Trainee		•		•		Others [59]
Dis	Exp	MPP	Rectification	RT	Trainee		•		•		Others [59]

provide suggestions for improvement [9, 47]. *Graph* was instead used in four setups found in two papers. These depict feedforward motion sequences over time along the horizontal axis, with data points showing trainee errors at corresponding moments of posture completions [16, 34]. As such, this form of feedback typically provides the *magnitude* or the error. Both *text* and *graph* feedback are all found to be presented as floating panels in the *environment*, and none have attempted to explore alternate ways to place this feedback, such as on the *trainee's body*, although situated visualization has revealed this possibility [37].

Detection feedback. In the 35 setups that incorporated ADDITIONAL corrective feedback, 11 of them employed feedback at *detection*-level, and they all presented it in the form of *color*. Among these 11 cases, two [7, 40] presented feedback *post-hoc* in the *environment*, while the remaining nine changed the *color* of *feedforward* in *real-time*. This preference towards the latter is likely because *detection*-level feedback provides only binary information, which is usually presented by an abrupt change in visual cue. In contrast to *direction* and *size*, the abrupt change in *color* does not obscure the *feedforward* and would therefore feel more natural, as they are commonly encountered in daily life (e.g., traffic lights).

Magnitude feedback. Seven setups provided *magnitude* feedback, and 4 of them used *graph*, while 3 utilized *color* hue to represent the deviation distance.

Rectification feedback. 17 setups featured *rectification*-level feedback. Two papers [9, 47] presented this feedback using *text*, whereas the one by Oshita et al. [47] used both *text* and *direction* to indicate where the limb should be moved to. Apart from this, 14 other setups used *direction*. This preference is understandable as directional cues can clearly guide the correction of body movements. A notable outlier is LightGuide [51], which used variations in the *area* of regions with different *colors* to instruct where to move the hand. Moreover, *text* can also supplement this cue by providing suggestions on the strength and speed of movement as well [9].

Size feedback. The occurrence of *size* is highly dependent on *direction*, which is the reason why feedback incorporating *size* is categorized under *rectification* rather than *magnitude*. Among the 13 setups that employed *size*, 12 represented it as the *length* of a *directional cue*; while the last one, LightGuide [51], used variations in the *area* of different colors to indicate the *rectification*.

Potential opportunities for visual features. In Section 5.4, we described for INTUITIVE feedback three features (*position*, *orientation* and *shape*); and for the **encoding** of ADDITIONAL feedback three features (*color*, *direction* and *size*) and two cues (*text* and *graph*). While we based our design space on both the literature and our own experience, there exist unused visual pre-attentive features, and novel uses of previously used features. For example, when utilizing a 3D arrow pointing at the target position to correct a movement error, we can provide additional *magnitude* feedback through its *rotation speed* (unused feature) and *shape* (used only for INTUITIVE feedback), i.e., the greater the *magnitude* of the error, the *faster* it rotates and the greater the *curvature* of the arrow.

7 USING THE DESIGN SPACE

We now briefly illustrate how our design space can be used to generate and describe new XR-based motion guidance systems based on two hypothetical scenarios: learning sign language and practicing deadlifts.

7.1 Scenario 1: Sign Language

Suppose the British Sign Language (BSL) alphabet is to be taught using XR. As the BSL alphabet comprises a unique hand pose for each letter, the feedforward should be *explicit* to allow the trainee to simply recognize and imitate it. A *discrete* update strategy can help ensure trainees are given enough time to accurately look at and perform each hand pose [26]. In terms of perspective, the feedforward may either be *MPP* to imitate traditional sign language training with an instructor, or be *IPP* as BSL movements occur entirely in the frontal field-of-view of the trainee, or even a combination of both *MPP* and *IPP* for redundancy.

In this approach, INTUITIVE feedback is already built into the *explicit* feedforward. However, ADDITIONAL feedback may assist the trainee further, especially for complete beginners. This may include the use of *color* to provide *detection* feedback of misaligned hand and finger joints, or possibly *directional arrows* to guide trainees to *rectify* said misalignments. Whilst *text* could be employed, providing feedback for each individual joint similar to Figure 4d may be excessive. An alternative would be to instead provide feedback for the whole posture, such as a binary *detection* that states whether the pose has been achieved or not, or a text description explaining how to *rectify* the mistake [9]. Lastly, *graph* would not be useful in this scenario as it mainly provides time-varying scalar data, which is not relevant for *discrete* postural guidance.

7.2 Scenario 2: Deadlifting

Now suppose XR is used to learn and perform deadlifts, a strength training exercise wherein the trainee is required to lift a barbell from the ground starting from a bent position until they are standing upright, all while maintaining a straight back. This motion involves changes in angles at multiple joints of the body and is vertically symmetrical. Thus, in addition to using *MPP* similar to a physical mirror found in a gym, *3PP* can also present the side-view of the trainee to keep track of body angles in real time—especially of their back. Because deadlifts are a physically intensive exercise, *explicit* feedforward can be used to avoid needless cognitive load. For the same reason, an *autonomous* update strategy can be used, as both *continuous* and *discrete* strategies impose posture checkpoints which may unnecessarily shift the trainee's focus towards satisfying said checkpoints, rather than safely performing the full range of motion in one smooth movement.

This type of training is reliant on the coordination of full-body movements throughout the entire motion. Therefore, in addition to the use of *color* and *arrows* as in Scenario 1, it is also possible to present changes in the key angles of the body which the trainee can study *post-hoc*. This allows for the safe analysis of mistakes whilst not under the physical stress of the exercise. A *graph* can be used to present this information, possibly with an accompanying *text* which provides more actionable suggestions that a personal trainer might typically give (e.g., “Your waist extended too fast!”).

8 DISCUSSION

During our review of the literature, we observed other factors that tend to be considered in addition to the aforementioned feedforward and feedback design. These factors are mainly contextual considerations and/or constraints that influence the design of the XR-based motion guidance system as a whole. In this section, we discuss these factors and provide guidelines and implications for future designers and researchers in XR-based motion guidance.

8.1 Motion Features

A motion sequence can be defined by several features that comprise the movement and determine how it is to be performed. We briefly describe a non-exhaustive list of the main motion features we found in the literature.

Speed. Certain motion types and training purposes may consider speed as important. For instance, ballroom dancing requires mastery of both footwork and timing [21], whereas most rehabilitation exercises might only require postural accuracy [53]. Motions involving speed would need to indicate this temporal aspect either *autonomously* or through some *additional contextual cue*. However, we note that the design of the motion feedforward would likely influence the trainee's ability to meet the speed requirements. For instance, Elsayed et al. [15] found that first-person perspectives had a higher angular error in fast movements, but third-person perspectives instead had a lower angular error for these fast movements.

Involved Limbs. As more limbs are involved in the motion, the demands for body coordination increase, thus making it harder to execute [15]. Not only does the trainee need to pay attention more to their limbs, but the overall spatial area around the trainee which the motion is performed in naturally increases. This, in turn, reduces the overall space that can be dedicated for motion feedforward and feedback, as most of it would need to be allocated to the trainee performing the motion itself.

Motion Area. This refers to the spatial extent and areas in which the motion is executed. This is directly related to the manner which the feedforward is presented: the feedforward should ideally be presented in such a manner that it encompasses the entirety of the motion area within the trainee's field-of-view. As a practical example, consider motions that require both arms to be along the left and right periphery of the trainee's field-of-view [59]. While the *first-person perspective* requires head turns caused in looking at the feedforward for either arm, the *third-person perspective* allows both to be viewed simultaneously. Such *third-person perspectives* have been found to perform significantly better for motions performed *behind* the trainee as well [15].

Trajectories. Many physical motions can be considered a form of spatiotemporal trajectory. As such, many motion guidance systems rely on showing the paths that body joints should follow in space [52, 53]. Yu et al. [59] summarized different features of motion trajectories that could influence the complexity of motions: trajectory length, number of intersections, number of arm joints, and the number of 2D segments. These features may induce physical fatigue, limb coordination challenges, and cognitive load for the trainees. While they did not conduct further analysis on these

features, highly complex motions would likely require more carefully designed feedforward and feedback representations, taking into account challenges such as visual clutter and overlapping or crossing trajectories.

8.2 Objective Scoring Framework

The objective scoring framework is used to provide overall assessment feedback to trainees with respect to postural [15, 39, 53], temporal [39, 51, 58], or other task-based performance such as basketball shooting accuracy [40] and punching speed [9]. It is determined by the training purpose and constrains decisions regarding feedforward visualization. For instance, when there is a requirement for trainees' *temporal* performance, it will be necessary for the feedforward to provide velocity visualization, demonstrating the temporal aspects of the motion to the trainee [58]. In the case of *postural* aspect, if correctness is calculated only at the key postures, then the **level of indirection** of the motion feedforward should contain a *discrete* component, to emphasize the significance of specific postures. This is due to the fact that trainees may lose patience and focus during a long motion sequence. However, the constraint imposed by the objective scoring framework on feedforward design is unidirectional. For instance, when feedforward is updated in a *discrete* manner, both key-posture-based [9] and frame-by-frame calculations [15, 59] can be found in the literature.

8.3 Individuality

While we refer to a "trainee" as a hypothetical user, in reality, everyone has unique physiological traits, cognitive abilities, learning styles, and emotional factors that directly impact their progress in motor learning. For instance, Li et al. [38] found that individuality affects image perception. Although not related to motion guidance systems, their work does suggest that varying levels of visual perception between people could influence the effectiveness of the motion guidance system—in this case, one that is reliant on vision. Moreover, the motion guidance system could be tailored to support the trainee's goals, such as by adjusting the difficulty of motions presented to them, or to support their current familiarity with the presented motion, such as by varying the thresholds that determine the **information level** of corrective feedback, or by increasing the **level of indirection** to increase the cognitive need for recall of the tasked motion [39].

8.4 Similarity in Appearance

As previously alluded to, the choice of technology (i.e., AR vs VR) influences the appearance of the trainee. In the case of AR, it is natural to default to either a human-like avatar, or to just rely on the trainee's own body. In the case of VR, avatars typically take on a wide array of forms, from the same human-like avatars [14], to more abstract looking ball-and-stick models of skeletal limbs and joints [26], to bodies formed out of capsules [47]. Previous work by Dürr et al. [14] has indicated that using humanoid avatars for the trainee can evoke higher performance regarding movement accuracy as compared to more abstract avatars. There appears however to be a limited body of research addressing a broader spectrum of similarity in appearance between explicit posture guidance and avatars. Given that motion guidance tasks inherently encompass

elements related to visual search and matching, there is a plausible basis for asserting that the similarity in appearance between avatars and feedforward cues can influence trainees' performance.

9 CONCLUSION

In this paper, we investigated the visual design of motion feedforward and corrective feedback in XR-based motion guidance system. We first surveyed the literature to understand how researchers have designed such systems and formed our corpus. We proposed four dimensions each for motion feedforward and corrective feedback, then discussed the interplay between feedforward and feedback based on our corpus. We also showed how to use our design space to create new systems using two hypothetical scenarios as examples. We then discussed the additional considerations that stem from the context of the motion guidance.

While outside of the scope of our literature survey, we note that several motion guidance systems consider additional forms of data that could further enhance or improve training effectiveness and outcomes. For instance, biometric sensor information is widely used in rehabilitation guidance (e.g., electromyographic [43] or electrical impedance tomography [62]). As XR technologies develop, it is likely that such data becomes readily accessible to use in any motion guidance system. For instance, eye-tracking may help improve the visibility of both feedforward and feedback cues during complex motions with large *motion areas*, whereas heart rate may support more *individuality* between trainees with different fitness levels.

Finally, as discussed in Section 8, there are numerous factors that influence the effectiveness of motion guidance systems. While we described how each factor may alter the design choices used with respect to our design space, their effect size and core mechanisms remain unexplored. Thus, the relative importance of each one remains unclear—especially when the context is considered. It may be that the *motion features* governs all other design choices in some contexts, whilst the *objective scoring framework* and the intended learning goal could be prioritized heavily in others. Moving forward, future work could examine how these priorities influence the design of motion guidance systems.

ACKNOWLEDGMENTS

This work is funded by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2075 – 390740016. We acknowledge the support by the Stuttgart Center for Simulation Science (SimTech).

REFERENCES

- [1] Fraser Anderson, Tovi Grossman, Justin Matejka, and George Fitzmaurice. 2013. YouMove: enhancing movement training with an augmented reality mirror. In *ACM Symp. User Interface Software and Technology (UIST)*. 311–320. <https://doi.org/10.1145/2501988.2502045>
- [2] Ricardo R Barioni, Willams Costa, Alessandra Aleluia, and Veronica Teichrieb. 2019. BalletVR: a Virtual Reality System for Ballet Arm Positions Training. In *Symp. Virtual and Augmented Reality (SVR)*. IEEE, 10–16. <https://doi.org/10.1109/SVR2019.00018>
- [3] Olivier Bau and Wendy E Mackay. 2008. OctoPocus: a dynamic guide for learning gesture-based command sets. In *ACM Symp. User Interface Software and Technology (UIST)*. 37–46. <https://doi.org/10.1145/1449715.1449724>
- [4] James William Burke, Michael McNeill, Darryl Charles, Philip Morrow, Jacqui Crosbie, and Suzanne McDonough. 2009. Serious games for upper limb rehabilitation following stroke. In *2009 Conference in Games and Virtual Worlds for Serious Applications*. IEEE, 103–110. <https://doi.org/10.1109/VS-GAMES.2009.17>
- [5] Benjamin Butz, Alexander Jussen, Asma Rafi, Gregor Lux, and Jens Gerken. 2022. A Taxonomy for Augmented and Mixed Reality Applications to Support Physical Exercises in Medical Rehabilitation—A Literature Review. In *Healthcare*, Vol. 10. MDPI, 646. <https://doi.org/10.3390/healthcare10040646>
- [6] Yuanzhi Cao, Xun Qian, Tianyi Wang, Rachel Lee, Ke Huo, and Karthik Ramani. 2020. An exploratory study of augmented reality presence for tutoring machine tasks. In *ACM Conf. Human Factors in Computing Systems (CHI)*. 1–13. <https://doi.org/10.1145/3313831.3376688>
- [7] Polona Caserman, Shule Liu, and Stefan Göbel. 2021. Full-Body Motion Recognition in Immersive-Virtual-Reality-Based Exergame. *IEEE Trans. Games* 14, 2 (2021), 243–252. <https://doi.org/10.1109/TG.2021.3064749>
- [8] Kaban Chacour and Georges Badr. 2016. Computer vision guidance system for indoor navigation of visually impaired people. In *IEEE Int. Conf. intelligent systems (IS)*. IEEE, 449–454. <https://doi.org/10.1109/IS.2016.7737460>
- [9] Hao Chen, Yujia Wang, and Wei Liang. 2022. VCoach: Enabling Personalized Boxing Training in Virtual Reality. In *IEEE Conf. Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 546–547. <https://doi.org/10.1109/VRW55335.2022.00124>
- [10] Alana Elza Fontes Da Gama, Thiago Menezes Chaves, Lucas Silva Figueiredo, Adriana Baltar, Ma Meng, Nassir Navab, Veronica Teichrieb, and Pascal Fallavollita. 2016. MirrARbilitation: A clinically-related gesture recognition interactive tool for an AR rehabilitation system. *Computer methods and programs in biomedicine* 135 (2016), 105–114. <https://doi.org/10.1016/j.cmpb.2016.07.014>
- [11] William Delamare, Céline Coutrix, and Laurence Nigay. 2015. Designing guiding systems for gesture-based interaction. In *Proc. ACM SIGCHI Symp. Engineering Interactive Computing SystemsJune (EICS)*. 44–53. <https://doi.org/10.1145/2774225.2774847>
- [12] William Delamare, Thomas Janssoone, Céline Coutrix, and Laurence Nigay. 2016. Designing 3D gesture guidance: visual feedback and feedforward design options. In *ACM Int. Working Conf. Advanced Visual Interfaces (AVI)*. 152–159. <https://doi.org/10.1145/2909132.2909260>
- [13] Florian Diller, Gerik Scheuermann, and Alexander Wiebel. 2022. Visual Cue Based Corrective Feedback for Motor Skill Training in Mixed Reality: A Survey. *IEEE Trans. Visualization and Computer Graphics (TVCG)* (2022). <https://doi.org/10.1109/TVCG.2022.3227999>
- [14] Maximilian Dürr, Rebecca Weber, Ulrike Pfeil, and Harald Reiterer. 2020. EGuide: Investigating different visual appearances and guidance techniques for egocentric guidance visualizations. In *Proc. Int. Conf. Tangible, Embedded, and Embodied Interaction (TEI)*. 311–322. <https://doi.org/10.1145/3374920.3374945>
- [15] 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 *ACM Symp. Virtual Reality Software and Technology (VRST)*. 1–10. <https://doi.org/10.1145/3562939.3565635>
- [16] Felix Escalona, Ester Martinez-Martin, Edmanuel Cruz, Miguel Cazorla, and Francisco Gomez-Donoso. 2020. EVA: EVAluating at-home rehabilitation exercises using augmented reality and low-cost sensors. *Virtual Reality* 24 (2020), 567–581. <https://doi.org/10.1007/s10055-019-00419-4>
- [17] Katherine Kennedy, Jeremy Hartmann, Quentin Roy, Simon Tangi Perrault, and Daniel Vogel. 2021. Octopocus in VR: using a dynamic guide for 3D mid-air gestures in virtual reality. *IEEE Trans. Visualization and Computer Graphics (TVCG)* 27, 12 (2021), 4425–4438. <https://doi.org/10.1109/TVCG.2021.3101854>
- [18] Dustin Freeman, Hrvoje Benko, Meredith Ringel Morris, and Daniel Wigdor. 2009. ShadowGuides: visualizations for in-situ learning of multi-touch and whole-hand gestures. In *Proc. the ACM Int. Conf. interactive tabletops and surfaces (ITS)*. 165–172. <https://doi.org/10.1145/1731903.1731935>
- [19] Shiro Furukawa, Toshihiro Uchibayashi, Toru Abe, and Takuo Suganuma. 2018. Design of an AR-supported System for Skill Training. In *2018 Tohoku-Section Joint Convention*. 213–213.
- [20] Patrick Gebhardt, Maximilian Weiß, Pascal Huszár, Xingyao Yu, Alexander Achberger, Xiaobing Zhang, and Michael Sedlmaier. 2023. Auxiliary Means to Improve Motion Guidance Memorability in Extended Reality. In *IEEE Conf. Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 689–690. <https://doi.org/10.1109/VRW58643.2023.00187>
- [21] Sebastian Fernando Chinchilla Gutierrez, Jose Victorio Salazar Luces, and Yasuhisa Hirata. 2022. Modality Influence on the Motor Learning of Ballroom Dance with a Mixed-Reality Human-Machine Interface. In *IEEE/SICE Int. Symp. System Integration (SII)*. IEEE, 177–182. <https://doi.org/10.1109/SII52469.2022.9708803>
- [22] 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 *Proc. Augmented Human International Conference (AH)*. 1–4. <https://doi.org/10.1145/2875194.2875237>
- [23] Ping-Hsuan Han, Yang-Sheng Chen, Yilun Zhong, Han-Lei Wang, and Yi-Ping Hung. 2017. My Tai-Chi coaches: an augmented-learning tool for practicing Tai-Chi Chuan. In *Proc. Augmented Human International Conference (AH)*. 1–4. <https://doi.org/10.1145/3041164.3041194>
- [24] Christopher Healey and James Enns. 2011. Attention and visual memory in visualization and computer graphics. *IEEE Trans. Visualization and Computer Graphics (TVCG)* 18, 7 (2011), 1170–1188. <https://doi.org/10.1109/TVCG.2011.127>

- [25] Christopher G Healey and James T Enns. 1999. Large datasets at a glance: Combining textures and colors in scientific visualization. *IEEE Trans. Visualization and Computer Graphics (TVCG)* 5, 2 (1999), 145–167.
- [26] Thuong N Hoang, Martin Reinoso, Frank Vetere, and Egemen Tanin. 2016. One-body: remote posture guidance system using first person view in virtual environment. In *Proc. Nordic Conf. Human-Computer Interaction (NordiCHI)*. 1–10. <https://doi.org/10.1145/2971485.2971521>
- [27] 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), 43. <https://doi.org/10.3389/frobt.2019.00043>
- [28] Atsuki Ikeda, Dong-Hyun Hwang, and Hideki Koike. 2019. A real-time projection system for golf training using virtual shadow. In *IEEE Conf. Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 1527–1528. <https://doi.org/10.1109/VRW19.8798196>
- [29] Atsuki Ikeda, Dong-Hyun Hwang, Hideki Koike, Gerd Bruder, Shunsuke Yoshimoto, and Sue Cobb. 2018. AR based Self-sports Learning System using Decayed Dynamic TimeWarping Algorithm.. In *ICAT-EGVE 2018 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*. 171–174. <https://doi.org/10.2312/egve.20181330>
- [30] Yeonsub Jin, Changmok Chun, Yeonhwa Cha, Gi-Hun Yang, and Sungchul Kang. 2013. Vibrotactile Cues for Motion Guidance. In *Intelligent Autonomous Systems 12: Volume 2 Proceedings of the 12th International Conference IAS-12, held June 26-29, 2012, Jeju Island, Korea*. Springer, 651–659. https://doi.org/10.1007/978-3-642-33932-5_61
- [31] Marcelo Kallmann, Carlo Camporesi, and Jay Han. 2015. Vr-assisted physical rehabilitation: Adapting to the needs of therapists and patients. In *Virtual Realities: International Dagstuhl Seminar, Dagstuhl Castle, Germany, June 9-14, 2013, Revised Selected Papers*. Springer, 147–168. https://doi.org/10.1007/978-3-319-17043-5_9
- [32] Daiki Kodama, Takato Mizuho, Yuji Hatada, Takuji Narumi, and Michitaka Hirose. 2023. Effects of Collaborative Training Using Virtual Co-embodiment on Motor Skill Learning. *IEEE Trans. Visualization and Computer Graphics (TVCG)* 29, 5 (2023), 2304–2314. <https://doi.org/10.1109/TVCG.2023.3247112>
- [33] Felix Kosmalla, Florian Daiber, Frederik Wiehr, and Antonia Krüger. 2017. Climbingvis: Investigating in-situ visualizations for understanding climbing movements by demonstration. In *Proc. ACM Int. Conf. Interactive Surfaces and Spaces (ISS)*. 270–279. <https://doi.org/10.1145/3132272.3134119>
- [34] Matthew Kyan, Guoyu Sun, Haiyan Li, Ling Zhong, Paisarn Munesawang, Nan Dong, Bruce Elder, and Ling Guan. 2015. An approach to ballet dance training through ms kinect and visualization in a cave virtual reality environment. *ACM Tran. Intelligent Systems and Technology (TIST)* 6, 2 (2015), 1–37. <https://doi.org/10.1145/2735951>
- [35] Thibaut Le Naour, Ludovic Hamon, and Jean-Pierre Bresciani. 2019. Superimposing 3D virtual self+ expert modeling for motor learning: Application to the throw in American football. *Frontiers in ICT* 6 (2019), 16. <https://doi.org/10.3389/fict.2019.00016>
- [36] Benjamin Lee, Dave Brown, Bongshin Lee, Christophe Hurter, Steven Drucker, and Tim Dwyer. 2020. Data Visceralization: Enabling deeper understanding of data using virtual reality. *IEEE Trans. Visualization and Computer Graphics (TVCG)* 27, 2 (2020), 1095–1105. <https://doi.org/10.1109/TVCG.2020.3030435>
- [37] Benjamin Lee, Michael Sedlmaier, and Dieter Schmalstieg. 2023. Design Patterns for Situated Visualization in Augmented Reality. *IEEE Trans. Visualization and Computer Graphics (TVCG)* (2023). <https://doi.org/10.1109/TVCG.2023.3327398>
- [38] Aoqi Li and Zhenzhong Chen. 2018. Personalized visual saliency: Individuality affects image perception. *IEEE Access* 6 (2018), 16099–16109. <https://doi.org/10.1109/ACCESS.2018.2800294>
- [39] Klemen Lilija, Søren Kyllingsbæk, and Kasper Hornbæk. 2021. Correction of avatar hand movements supports learning of a motor skill. In *IEEE Conf. Virtual Reality and 3D User Interfaces (VR)*. IEEE, 1–8. <https://doi.org/10.1109/VR50410.2021.00069>
- [40] Tica Lin, Rishi Singh, Yalong Yang, Carolina Nobre, Johanna Beyer, Maurice A Smith, and Hanspeter Pfister. 2021. Towards an understanding of situated ar visualization for basketball free-throw training. In *ACM Conf. Human Factors in Computing Systems (CHI)*. 1–13. <https://doi.org/10.1145/3411764.3445649>
- [41] Alejandro Martin-Gomez, Haowei Li, Tianyu Song, Sheng Yang, Guangzhi Wang, Hui Ding, Nassir Navab, Zhe Zhao, and Mehran Armand. 2023. STTAR: surgical tool tracking using off-the-shelf augmented reality head-mounted displays. *IEEE Trans. Visualization and Computer Graphics (TVCG)* (2023). <https://doi.org/10.1109/TVCG.2023.3238309>
- [42] Neil McHenry, Tanner Hunt, William Young, Anthony Gardner, Ujwala Bhagavatula, Blake Bontz, Jonathan Chiù, Gregory Chamitoff, and Ana Diaz-Artiles. 2020. Evaluation of Pre-Flight and On Orbit Training Methods Utilizing Virtual Reality. In *Proceedings of the AIAA Scitech 2020 Forum*. <https://doi.org/10.2514/6.2020-0168>
- [43] Marina Melero, Annie Hou, Emily Cheng, Amogh Tayade, Sing Chun Lee, Mathias Unberath, Nassir Navab, et al. 2019. Upbeat: augmented reality-guided dancing for prosthetic rehabilitation of upper limb amputees. *Journal of healthcare engineering* 2019 (2019). <https://doi.org/10.1155/2019/2163705>
- [44] Fariba Mostajeran, Nikolaos Katzakis, Oscar Ariza, Jann Philipp Freiwald, and Frank Steinicke. 2019. Welcoming a holographic virtual coach for balance training at home: two focus groups with older adults. In *IEEE Conf. Virtual Reality and 3D User Interfaces (VR)*. IEEE, 1465–1470. <https://doi.org/10.1109/VR.2019.8797813>
- [45] Tamara Munzner. 2014. *Visualization analysis and design*. CRC press.
- [46] Andreea Muresan, Jess McIntosh, and Kasper Hornbæk. 2023. Using Feedforward to Reveal Interaction Possibilities in Virtual Reality. *ACM Trans. Computer-Human Interaction* (2023). <https://doi.org/10.1145/3603623>
- [47] Masaki Oshita, Takumi Inao, Tomohiko Mukai, and Shigeru Kuriyama. 2018. Self-training system for tennis shots with motion feature assessment and visualization. In *IEEE Int. Conf. Cyberworlds (CW)*. IEEE, 82–89. <https://doi.org/10.1109/CW.2018.00025>
- [48] Washington X Quevedo, Jessica S Ortiz, Paola M Velasco, Jorge Sánchez, Marcelo Álvarez V, David Rivas, and Víctor H Andaluz. 2017. Assistance system for rehabilitation and valuation of motor skills. In *Augmented Reality, Virtual Reality, and Computer Graphics: 4th International Conference, AVR 2017, Ugento, Italy, June 12–15, 2017, Proceedings, Part II* 4. Springer, 166–174. https://doi.org/10.1007/978-3-319-60928-7_14
- [49] Christian Schönauer, Kenichiro Fukushi, Alex Olwal, Hannes Kaufmann, and Ramesh Raskar. 2012. Multimodal motion guidance: techniques for adaptive and dynamic feedback. In *Proc. ACM Int. Conf. Multimodal interaction (ICMI)*. 133–140. <https://doi.org/10.1145/2388676.2388706>
- [50] Yoones A Sekhavat and Mohammad S Namani. 2018. Projection-based AR: Effective visual feedback in gait rehabilitation. *IEEE Tran. Human-Machine Systems* 48, 6 (2018), 626–636. <https://doi.org/10.1109/THMS.2018.2860579>
- [51] Rajinder Sodhi, Hrvoje Benko, and Andrew Wilson. 2012. LightGuide: projected visualizations for hand movement guidance. In *ACM Conf. Human Factors in Computing Systems (CHI)*. 179–188. <https://doi.org/10.1145/2207676.2207702>
- [52] Mauricio Sousa, João Vieira, Daniel Medeiros, Artur Arsenio, and Joaquim Jorge. 2016. SleeveAR: Augmented reality for rehabilitation using realtime feedback. In *ACM Int. Conf. Intelligent User Interfaces (IUI)*. 175–185. <https://doi.org/10.1145/2856767.2856773>
- [53] 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 *ACM Conf. Human Factors in Computing Systems (CHI)*. 4123–4132. <https://doi.org/10.1145/2702123.2702401>
- [54] Lain Turmo Vidal, Hui Zhu, and Abraham Riego-Delgado. 2020. Bodylights: Open-ended augmented feedback to support training towards a correct exercise execution. In *ACM Conf. Human Factors in Computing Systems (CHI)*. 1–14. <https://doi.org/10.1145/3313831.3376268>
- [55] Thomas Waltemate, Irene Senna, Felix Hülsmann, Marieke Rohde, Stefan Kopp, Marc Ernst, and Mario Botsch. 2016. The impact of latency on perceptual judgments and motor performance in closed-loop interaction in virtual reality. In *ACM Symp. Virtual Reality Software and Technology (VRST)*. 27–35. <https://doi.org/10.1145/2993369.2993381>
- [56] Liu Wang, Mengjie Huang, Rui Yang, Hai-Ning Liang, Ji Han, and Ying Sun. 2022. Survey of movement reproduction in immersive virtual rehabilitation. *IEEE Trans. Visualization and Computer Graphics (TVCG)* (2022). <https://doi.org/10.1109/TVCG.2022.3142198>
- [57] Claes Wohlin. 2014. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In *Proc. Int. conf. evaluation and assessment in software engineering (EASE)*. 1–10. <https://doi.org/10.1145/2601248.2601268>
- [58] Ungyeon Yang and Gerard Joungyun Kim. 2002. Implementation and evaluation of “just follow me”: An immersive, VR-based, motion-training system. *Presence* 11, 3 (2002), 304–323. <https://doi.org/10.1162/105474602317473240>
- [59] Xingyao Yu, Katrin Angerbauer, Peter Mohr, Deniz Kalkofen, and Michael Sedlmaier. 2020. Perspective matters: Design implications for motion guidance in mixed reality. In *IEEE Symp. Mixed and Augmented Reality (ISMAR)*. IEEE, 577–587. <https://doi.org/10.1109/ISMAR50242.2020.00085>
- [60] Xuan Zhang, Erwin Wu, and Hideki Koike. 2021. Watch-Your-Skiing: Visualizations for VR Skiing using Real-time Body Tracking. In *IEEE Symp. Mixed and Augmented Reality (ISMAR)*. IEEE, 387–388. <https://doi.org/10.1109/ISMAR-Adjunct54149.2021.00088>
- [61] Qiushi Zhou, Andrew Irlitti, Difeng Yu, Jorge Goncalves, and Eduardo Velloso. 2022. Movement guidance using a mixed reality mirror. In *ACM SIGCHI Conf. Designing Interactive Systems (DIS)*. 821–834. <https://doi.org/10.1145/3532106.3533466>
- [62] Junyi Zhu, Yuxuan Lei, Aashini Shah, Gila Schein, Hamid Ghaednia, Joseph Schwab, Casper Harteveld, and Stefanie Mueller. 2022. MuscleRehab: Improving Unsupervised Physical Rehabilitation by Monitoring and Visualizing Muscle Engagement. In *ACM Symp. User Interface Software and Technology (UIST)*. 1–14. <https://doi.org/10.1145/3526113.3545705>