Investigating Visual Perspectives on Guidance Visualisations for Motor Learning

Stefan Paul Feyer
HCI Group, University of Konstanz

Maximilian Dürr
Supervisor

Prof. Dr. Harald Reiterer
Examiner

Seminar to the Master's Project
December 2019

Abstract

With the improvements of Mixed Reality technology, motor learning in Mixed Reality has become a major topic in commercial companies and academic research. In the real world, guidance always takes place in the exo-centric perspective, but with Mixed Reality technology, ego-centric guidance becomes possible. While motor learning in mixed reality proofed to perform well, we know very little about the influence of the visual perspective on a guidance visualisation. This work focusses on the background of these topics with the aim to design a study which will be able to deliver data to investigate the influence of four different visual perspectives. This will lead to insights on which visual perspective to use for motor learning. The result of this work is a within-subject study design with four visual perspectives as conditions.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Motivation

In recent years, Mixed Reality (MR) devices became affordable[1], portable[2] and usable in many conditions. Not only academic researchers are interested in this technology, but commercial companies also found MR devices helpful to explore new possibilities to use them profitably. With this development, learning and training in MR became possible for many cases, too. EON[3], for example, calls themself "the world leader in Virtual Reality based knowledge transfer for industry, education, and edutainment". They develop MR programs for several platforms, e.g. intending to guide workers, reducing mistakes and thus reducing costs. These programs address a lot of use cases in the field of education, energy, health & medical, manufacturing & industrial, defence & security and aerospace. Tasks include e.g. ground crew training for a Boeing 777, augmented reality (AR) assembly training, exploring or anatomy simulation, to mention only a few.
Microsoft also stepped into this topic with partners, developing tools for apprenticeship, maintenance, or remote training. E.g. The Smart Glass experience Lab[4] of the Fraunhofer Institute use the MS Hololens[5] for remote maintenance, compare figure 1.1.
Motor learning plays a mayor role in our everyday life when it comes to – for example – sports, arts or dancing. Gaining proficiency in one of these areas is not possible without intense motor learning. Mostly a teacher or trainer is required for progress. For example, a student wants to learn a movement from a teacher. In the real world, the teacher stands in front of the student performing the movement and the student tries to mimic it. This visual perspective is called exo-centric or 3rd person perspective. If for example, a teacher is not available or affordable, other sources for motor learning are available like videos. Still, these videos are in the

[1]vive.com, oculus.comaccessed: 3.12.2019
[2]arvr.google.com/daydream accessed: 3.12.2019
[3]https://www.eonreality.com/ accessed: 14.12.2018
[4]fit.fraunhofer.de/de/fb/cscw/smart-glasses-experience-lab.html accessed: 18.11.2019
[5]microsoft.com/en-us/hololens accessed: 3.12.2019

Figure 1.1: Remote maintenance with Hololens by Smart Glass Experience Lab. The on-site worker wears a Hololens, while the remote trainer draws green hints to resolve a miss-wiring, taken from https://www.youtube.com/watch?v=1QFMPo5k6p0, accessed (30.07.2019)

exo-centric perspective. But, with today's MR technology we have the possibility to change this perspective what we cannot do in the real world. A student can "step into" the teacher's virtual body and see the instruction from the 1st person perspective of the teacher, also called ego-centric perspective. For developing MR learning and training environments, researchers put much effort into developing how-to's and guidelines to ensure proper systems e.g. [14]. However - as we will see in chapter 3 - there is a research gap about the visual perspective in these systems. The implication of the change of the visual perspective is not investigated sufficiently. This work aims to close this gap. This seminar thesis is the first out of three parts, followed by a master's project and a master's thesis. The overall aim of this work is to answer the following main research question:

MRQ Does the visual perspective on a virtual guidance visualisation have an influence on motor learning in MR environments.

The answer to this research question is important for designing motor learning systems for mixed reality.
The outcome of this work is a study design that will be able to address the research questions. In the master's project, the proposed study will be implemented. With this system a study will be conducted, generating the data to answer the research

question. The master's thesis itself will take the generated data to finally answer the research question.

1.2 Outline

For a well-developed study design, many aspects must be taken into consideration. The main aspects will be discussed in this work. Further aspects like algorithms and technology are discussed in the master's project. In chapter 2 this work sets the scope and provides theoretical foundations. in chapter 3 the parameters for the study design are discussed utilizing related work. With the scope and parameters set, chapter 4 proposes a study design which serves as the base for the master's project. In the end, an outlook on the master's project and master's thesis is given. Compare figure 1.2.

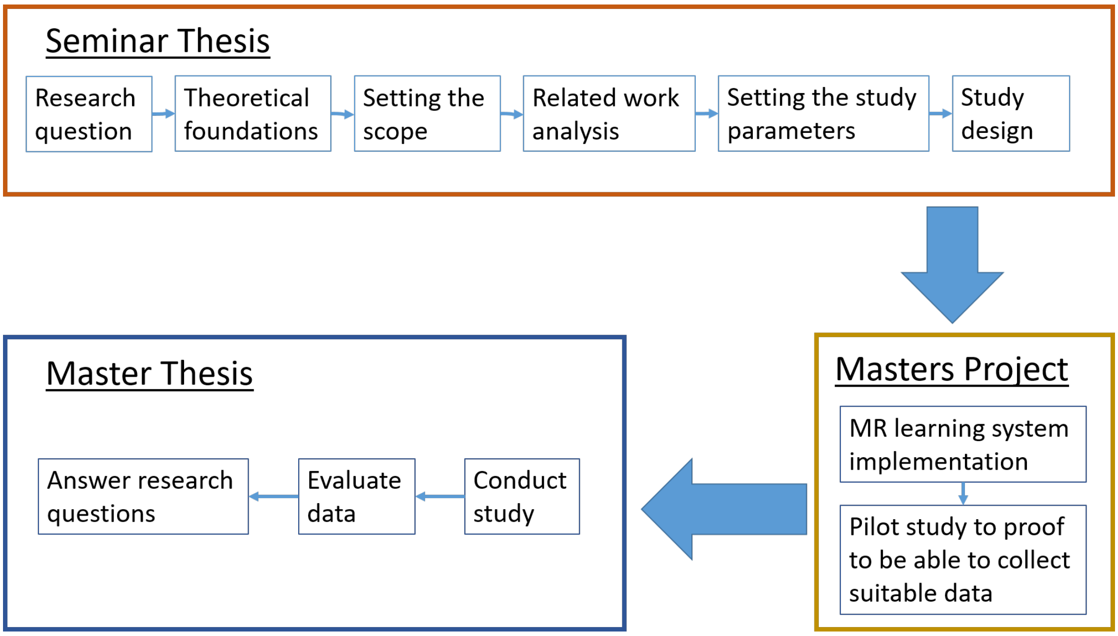


Figure 1.2: Overall process of the masters thesis.

2 Theory and Scope

In this chapter, the scope of the study will be set. We dig into visual perspectives and define them by a continuum from ego-centric to exo-centric. Movements will be classified and discussed how to measure them. Eventually, a definition of mixed reality is provided. For a well designed study, these theoretical foundations are vital. If not indicated otherwise, the content is adopted from the book Motor Learning and Skills [21].

2.1 Visual Perspectives

Wang and Milgram [26] describe the perspectives on the centricity continuum, see figure 2.1. On the most left-hand side of the continuum the ego-centric perspective is located. Ego-centric means that the anchor of the viewport is located inside the object to control - for simplicity, this object in question is referred to as an avatar. On the right-hand side the exo-centric perspective is located. This viewport anchor is a fixed point in the scene. The exo-centric perspective gives the user the possibility to examine the scene from a bird's-eye view. The movement or angle of the avatar has no influence on the viewports position or angle. The main difference is the so-called tether distance and the degree of freedom of the viewport. Milgram and Wang investigated on tethered viewports and define it as the distance between the eyes of the avatar and the viewport which is following the avatar. This describes the middle part of the continuum. Zero-distance viewport describes the ego-centric perspective. The longer the tether distance the more the perspective is located on the right of the continuum. They also distinguish between dynamic and rigid tethering relationships. A dynamic tethered viewport is controlled by the user in all six degrees of freedoms (DOF) while a rigid tethered viewport stands like a pole and can only be controlled in 3 DOF. Rigid tethered viewports are common in modern 3rd person computer games. In the next chapter, visual perspectives will be investigated in more detail.

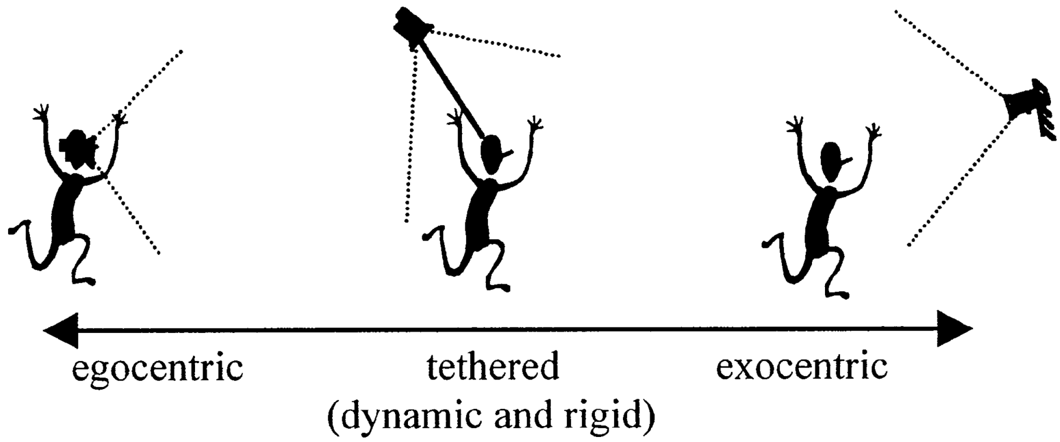


Figure 2.1: Centricity continuum by Wang and Milgram [26].

2.2 Motor Learning

Learning Movements

Motor learning is achieved through instruction, trying, imitation or a combination of them. Instructions can be written (see also figure 2.4 and further explanations in section 2.2), visual or verbal. Visual or verbal instruction include a trainer, teaching the student movements. In this case, verbal and physical feedback also plays a role in the learning process. The process of motor learning is divided into three parts. In the beginning of learning a technique, the student starts in the *cognitive* stage. In this stage, the students tries to figure out what to do to achieve the task. For this, high cognitive activity is required and strategies needs to be evaluated. The student's performance increases dramatically and is larger than in any other stage, but also inconsistent. The use of instructions and other training techniques are most effective. The next stage is called the *associative* stage. It begins when the student has determined the most effective way of doing the task. Performance increases more gradually but becomes more consistent. In the last *autonomous* stage, the performer gains proficiency and other simultaneous happening tasks are less likely to interfere.

> Since the use of training methods are most effective in the *cognitive* stage and the performance gain is highest in the *cognitive stage*, tasks in this stage are best suited for the study.

Movement Classification

Movements can be classified. There are two common classification schemas. The first one is based on the particular movements performed and is divided into *discrete*, *continuous* and *serial movements*, compare figure 2.2. The second one is based on perceptual attributes of the task and is divided into *open* and *closed skills*, compare figure 2.3. Both classification are represented by a continuum.

Discrete, Continuous and Serial Movements

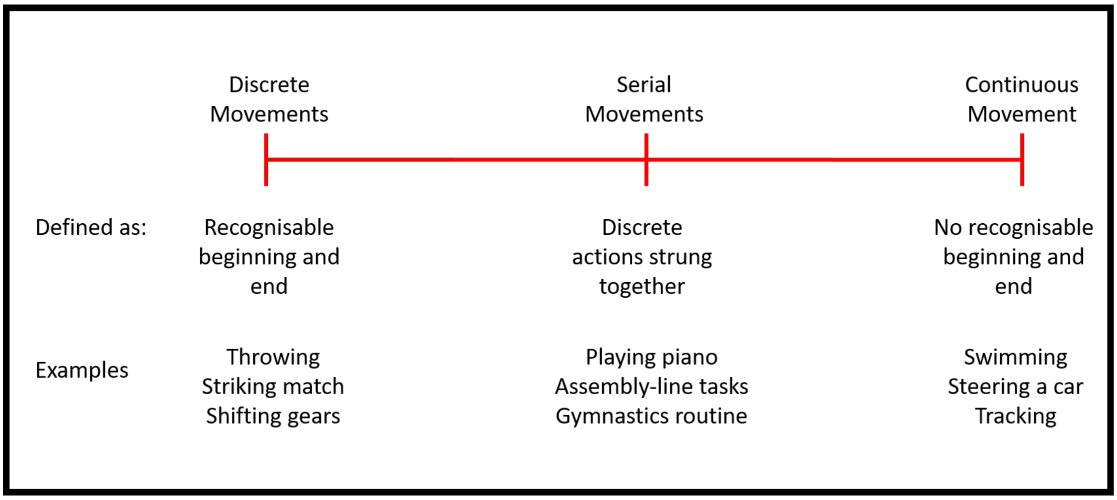


Figure 2.2: Continuum of movement classification based on the movement itself, redrawn from [21].

Discrete movements are located on the one end of the continuum. These are movements with a recognisable beginning and end. The end of a discrete movement is defined by the task itself and can be very rapid like blinking or longer like undersigning. Examples are kicking a ball, shifting gears in a car or striking a match.

Continuous movements are located on the other end of the continuum. These movements do not have a recognisable beginning or end. The behaviour continues till the movement arbitrarily stops. Continuous tasks tend to be longer than discrete tasks. Examples are swimming, running or steering a car.

Serial movements are located in the middle part of the continuum. Following the nature of a continuum, these movements are neither discrete nor continuous. They can consist of smaller movements tied together. Furthermore, serial movements can be rather long but are not stopped arbitrarily. Serial tasks can be seen as many discrete tasks strung together and the order (and sometimes timing) is important. Examples are starting a car or preparing and lighting a wood fireplace. The nature of *continuous movements* is having no recognizable beginning and end. This makes it hard to describe a distinctive task for a study design. While *discrete movements* are too short for a proper task, *serial movements* are most suitable for a study task.

Discrete movements are too short for a proper evaluation. Because of the nature of having no recognisable beginning and end, *Continuous movements,* also seems not to be suitable as a task in a study. Furthermore, in chapter 3 we will see that most researchers use *serial movements* as study task. Because of this, the scope is set for serial movements as study task.

Open and Closed Skills

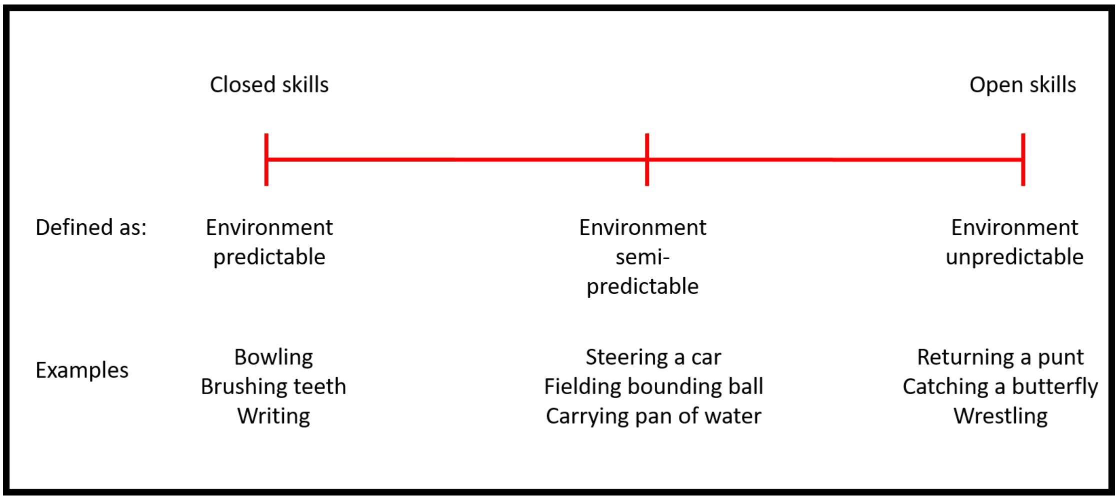


Figure 2.3: Continuum of movements classification based on perceptual attributes, redrawn from [21].

Open skills: The environment is unpredictably changing. The performer cannot plan his activity effectively in advance. Own movements depend on the environment. For example, if an ice hockey player shoots a puck, his movement dependents on the movement of the keeper. Another example is driving on a freeway. The driver needs to adjust his driving dependent on the behaviour of other cars. Success in open skills is largely determined by the extent to which an individual can adapt the planned motor behaviour to the changing environment.
Closed skills: The environment is predictable, mainly because it is stable. This

means that the performer can plan his activity. Examples are bowling, archery or singing. To evaluate only the motor learning and not environmental influences the study is conducted in a controlled environment in a laboratory. Thus only *closed skills* are taken into consideration.

This work aims to evaluate the influence of the visual perspective on motor learning, not how well students can react to unforeseen changes in the environment. Therefore, the scope is set to *closed skills*.

Measuring Movements

To measure a movement and matching it to a given motion is not a trivial task. Since e.g. dancing is a purely physical task, movements must be recognised, digitalised and judged. One approach is to use an analogue description for dancing and translate them into the digital world. Choensawat [5] began with Rudolph von Laban - a professional dancer. Von Laban developed a broadly used dance notation. His work lead to the *Laban Movement Analysis* with which human movements could be quantized[1]. There are four main components to systematically describe movements in the *Laban Movement Analysis*: body, effort, shape and space. Each component can describe movements independently or combined. Hachimura et al. [7] used the methodology of *Laban Movement Analysis* and adopted it for digital movements.

Yoshimura et al. [30] followed a similar approach from another dance movement description theory called *furi*. *Furi* is described by four so-called *indices*: *kamae*, *jyu-shin*, *koshi*, *uchiwa*. Yoshimura et al. map these indices to concrete markers on the body of a performer. Qian et al. [20] developed a gesture recognition system for performing arts. To match the motions ten body parts were defined: head, torso, upper arms, forearms, upper legs and lower legs. For each body part, the Mahalanobis distance is calculated to an ideal point. The Mahalanobis distance describes the distance between point p and distribution D.

To measure a movement, a common approach is to measure a single point on the subjects body repeatedly and then calculate an overall score, measuring how accurate the movement was. In literature, three main categories of those point measures are listed: *error of a single subject*, *measures of time and speed* and *measures of movement magnitude*.

[1] Brockhaus, Rudolf Laban. http://www.brockhaus.de/ecs/enzy/article/laban-rudolf (accessed 2018-10-25)

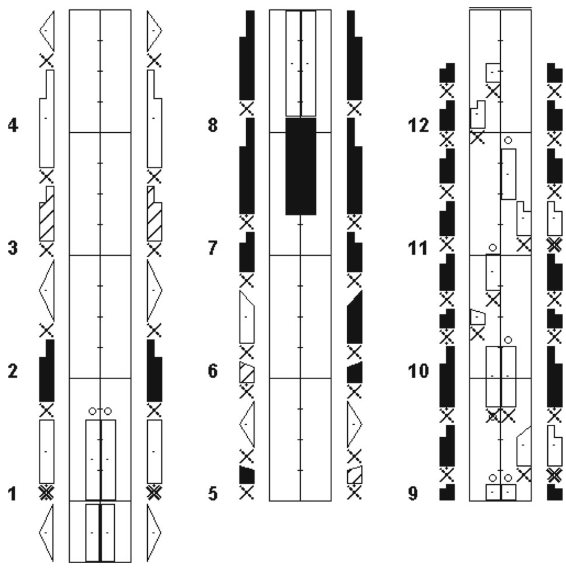


Figure 2.4: Laban notation. Generated through automatic movement interpretation by Choensawat [5].

> As chapter 3 will show, *error of a single object* and *measures of time and speed* is most commonly used and is most promising to generate the date in question. For this reason, these two measures will be used to evaluate the movements in the study proposed in chapter 4. *Measures of movement of magnitude* suit not for the evaluation, because of the nature of the task which is described in chapter 3.

Measures of Error for a Single Subject

Measures of error for a single subject represent the degree to which the target movement was missing. A target can be to perform an act at a particular time (timestamp), move with a certain force (amount of force) or hit a spatial target (a point in spatial volume). The attribute of the target serves as the variable in question. The error itself describes the distance - regarding the dimension - from the target. The following list gives an insight into the most important error measures.

- ***Constant Error*** describes the average error between the actual accuracy and the target. Means, in average the performer missed the target by CE.

$$CE = \frac{\sum_i (x_i - T)}{n} \tag{2.1}$$

with x_i: score, n: number of values, T: target value.

- ***Variable Error*** measures the inconsistency in movements. The more consistent the movements, the smaller VE. VE does not depend on whether or not the subject was close to the target.

$$VE = \sqrt{\frac{\sum(x_i - T)^2}{n}} \tag{2.2}$$

with x_i: score, n: number of values, T: target value.

- ***Total Variability*** describes the total variability around a target. The combination of VE and CE represents the total amount of spread around the target. It is an overall measure how successful the subject was in achieving the target.

$$E = VE^2 + CE^2 = \sqrt{\frac{\sum(x_i - T)^2}{n}} \tag{2.3}$$

with x_i: score, n: number of values, T: target value.

- ***Absolute Error*** is a measure of the overall accuracy in performance.

$$AE = \frac{\sum |x_i - T|}{n} \tag{2.4}$$

with x_i: score, n: number of values, T: target value.

- ***Absolute Constant Error*** is the absolute value of CE. Because of negative and positive values can cancel each other out

$$ACE = |CE| \tag{2.5}$$

Measures of Time and Speed

The basic idea of time and speed measures is, that a performer who can accomplish more in a given amount of time or who can accomplish a given amount of behaviours in less time is more skilful. Measures here are $\frac{time}{unit}$ or $\frac{units}{time}$.
The two most common examples are *reaction time* (RT) and *movement time* (MT). Reaction time describes the amount of time between a stimulus and the regarding start of a movement. This time span is important for two reasons. First, RT has high validity for real-life tasks. Secondly, RT measures the time taken for mental events like stimulus processing or decision making.
Movement time is the time interval between the end of the RT phase, meaning the start of the response, and the completion of the movement. The sum of RT and MT is called *response time*.

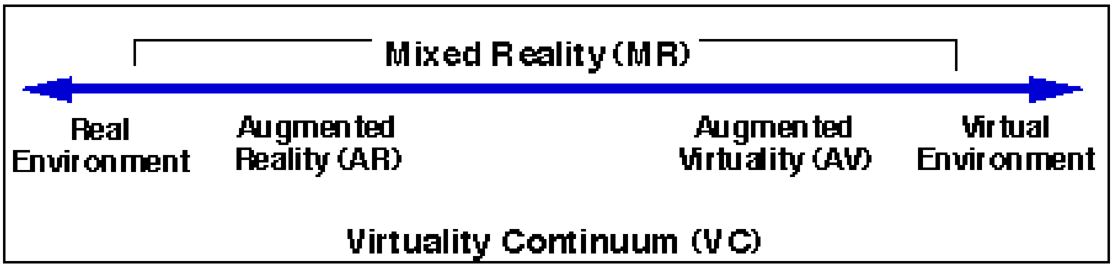


Figure 2.5: Mixed reality continuum by Milgram et al. [16].

Measures of Movement Magnitude

To measure skill, the produced magnitude of behaviour can be used. E.g. the distance a discus is thrown. A famous example is the "ski simulator". Rubber bands hold a plate centred between two poles. The magnitude, in this case, is the dislocation of the board from the centre by using full-body movements. As described, this measure can not be used for the task used in the proposed study design, thus only described in short.

2.3 Mixed Reality

Already in 1994 Milgram and Kishinho [16] defined the Mixed Reality continuum for displays, see figure 2.5. On the most left-hand side, the real world is located. On the right-hand side the purely digital world is located. The more we move from the left to the right, the more of the real world is replaced with digital elements. Respectively, starting from a purely digital environment and moving on the continuum to the left, the more of the digital environment is replaced by real-world elements. In virtual reality, the view on the real world is completely blocked. Common devices are e.g. the Oculus Rift or HTC Vive. Augmented reality can be realised by optical see-through devices like the Microsoft HoloLens. Seeing the own body during motor learning seems to be intuitively the best way to maintain the perception of their own body.

But for the following reason the scope is set to VR: AR technology is still limiting, especially because of their small field of view, which can influence guiding negatively [11]. Video see-through provides a larger field of view but also has limiting aspects like latency and distortion. In this case, the representation of the own body needs to be maintained by using motion capturing and rendering on the position of the own body.

2.4 Summary

This chapter provides theoretical foundations needed for the planned research in the field of perspectives and motor learning. Furthermore, the scope of the research was set. **Measures for single error** will be applied in the study. The task will be a **serial task** and also only **closed skills** will be trained. Eventually, the study will take place in **virtual reality**.

3 Related Work

Table 3.1 shows the previous work of researchers connected with motor learning in MR. In total there five classes of tasks they used for evaluation. The most common task is *arts*, like Martial Arts or Tai Chi [8–11, 13, 17, 18, 29]. But also dance movements [1–3, 7, 28], Sports [6, 12] and rehabilitation tasks [4, 19, 23–25] can be found in literature. Finally, some researchers decided to use abstract movements [15, 22] to evaluate their system like Anderson et al. [1] who also compared abstract with real-world movements.
The above literature can be also clustered into classes of guidance visualisations. The first class uses a person-shaped avatar performing a movement which the student can mimic (references: all mentions publications, except the publications mentioned in the second class). The second class uses variations of indicators like arrows or lines to guide the student [6, 11, 22–24]. Noticeable here is, that indicators are only used in scenarios where the guidance visualisation focusses on only a part of the body. For full-body movements indicators are too overwhelming because of providing too much attention points the student has to focus on [22].

Throughout the literature research became clear that five aspects of designing a

Ego-centric	**Exo-centric**	**Ego || Exo-centric**
AR-Arm (Han et al. 2016)	MotionMA (Velloso et al. 2013)	OneBody (Hoang et al. 2016)
Just Follow Me (Yan & Kim 2002)	YouMove (Anderson et al. 2013)	LightGuide (Sodhi et al. 2012)
Gohstman (Chinthammit et al. 2014)	VR Dance Trainer (Jacky Chan et al. 2010)	MR Dance Trainer (Hachimura et al. 2004)
Stylo and Handifact (Katzakis et al. 2017)	Physio@Home (Tang et al. 2015)	Free Throw Simulator (Covaci et al. 2014)
GhostHands (Scavo et al. 2015)	OutSide me (Yan et al. 2015)	Training Physical Skill (Kojima et al. 2014)
	e-Learning Martial Arts (Komura et al. 2006)	SleeveAR (Sousa et al. 2016)
	My Tai-Chi Coaches (Han et al. 2017)	Tai Chi Trainer (Chua et al. 2006)
	Performance Training (Chan et al. 2007)	
	RT Gestrue Recognition (Portillo et al. 2008)	
	KinoHaptics (Rajanna et al. 2015)	
	TIKL (Lieberman & Breazeal 2007)	

Table 3.1: Related work divided by visual perspectives.

	Tai Chi Trainer	YouMove	VR Dance Trainer	OneBody	LightGuide	Physio@Home
Perspective	Exo-centric, Ego & Augmented Exo-centric	Exo-centric	Exo-centric	Ego-centric, Exo-centric	Ego-centric, Exo-centric	Exo-centric
Task	Tai Chi	Dance (Ballet), abstract	Dance (HipHop)	Martial Arts	Abstract	Shoulder rehab
Guidance Visualisation	hr avatar, wireframe, mimic avatar	Stick figure, mimic avatar	hr figure, mimic avatar	Stick figure, mimic avatar	Indicators, follow/mimic	Indicators
Variables	Perspectives, performance measure	VR/Video, performance	Video/VR, performance	Training method, performance	Visualisations, Perspective, Performance	Visualisation, performance
Results	No difference in performance	VR better than video	VR better than video	Ego better than exo	Ego better than exo	Multi view better than single view

Table 3.2: Overview for the related work discussed in detail.

study for motor learning are important. First, how the perspective is implemented and what technology will be used. Second, what task is suitable for such a study. Third, what type of guidance visualisation to use. This section is structured in by three points of interest: how the guidance visualisation looks like, what degree of realism is suitable for a person-shaped avatar as guidance visualisation or if even feedback is necessary. Fourth, what measures to use and what to measure. And finally, five, what are the implication of the results of other studies for the study here in question. To analyse these five topics, we dive deeply into six works of other researchers, an overview of this works is given in table 3.1. In the next section, each of these topics are discussed by means of the related work. After each topic, a conclusion is drawn and the parameters for the study are set. With the scope and the parameters for the study in mind, the study design can be found in the next chapter.

3.1 Implemented Visual Perspective

	Tai Chi Trainer	YouMove	VR Dance Trainer	OneBody	LightGuide	Physio@Home
Perspective	Exo-centric, Ego & Augmented Exo-centric	Exo-centric	Exo-centric	Ego-centric, Exo-centric	Ego-centric, Exo-centric	Exo-centric

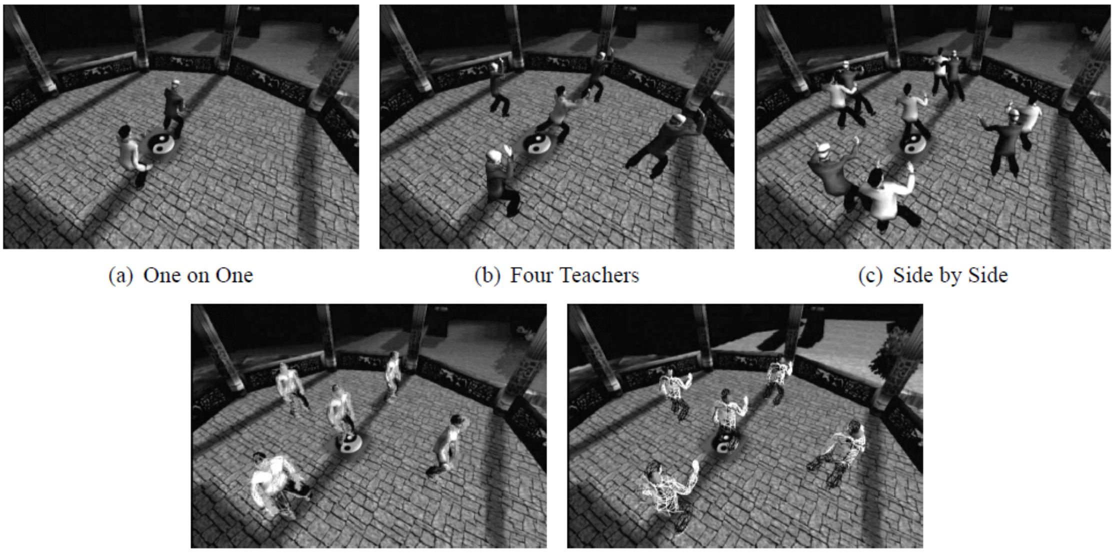
Table 3.3: Overview: implemented visual perspectives in the systems.

Chua et al. [17] implemented five different perspectives to teach with **Tai Chi Trainer**, compare figure 3.1. There are three exo-centric perspectives (figure 3.1a-c) and two ego-centric perspectives combined with exocentric (figure 3.1 d,e). In **One on One** the student stand next to one teacher, which is closest to a real world training scenario, see figure 3.1(a). In condition – **Four Teachers** – the student is surrounded by four teachers, with the student in the middle figure 3.1(b). The **Side by Side** figure 3.1(c) condition shows four pairs of teacher and student in the same formation as in condition (b). The first ego-centric & exo-centric combination is called **Superimposition 1** figure 3.1(d). Here the formation still remains, but the student in the middle is surrounded by four more students. On each of the student a red wireframe teacher is superimposed. **Superimposition 2** figure 3.1(e), is similar to Superimposition 1 and differs only in the visual representation: the student is now a green, transparent wireframe. Chua et al. use no pure ego-centric perspective, since in the latter two conditions the teacher can be seen also from the exo-centric perspective.

YouMove by Anderson et al. [1] uses Microsoft Kinect [1] to record motions and to track the student. The recorded instruction, as well as the student's movements, are projected on an AR mirror. The special about this mirror is, that the degree of reflection can be changed by simply adjusting the light in the room. With this, the student can see his own real body as a reflection as well as the projection of a beamer simultaneously. If the room is bright, the student sees only his reflection and in a dark room only the projection is visible.

YouMove provides multiple stages for the learning process: (1) Demonstration, (2) Posture Guide, (3) Movement Guide, (4) Mirror, (5) On your own. In stages 1-3, the student sees the guidance visualisation in the exo-centric perspective. In (1) the user sees a video of the movement. In stage 2 & 3 the user is superimposed by a stick figure. Stage 4 & 5 does not allow a view on the teacher's representation.

[1] https://developer.microsoft.com/de-de/windows/kinect, accessed 10.12.19



(a) One on One (b) Four Teachers (c) Side by Side

(d) Superimposition 1 (e) Superimposition 2

Figure 3.1: Implemented visual perspectives by Chua et al. [17] for Tai Chi Trainer.

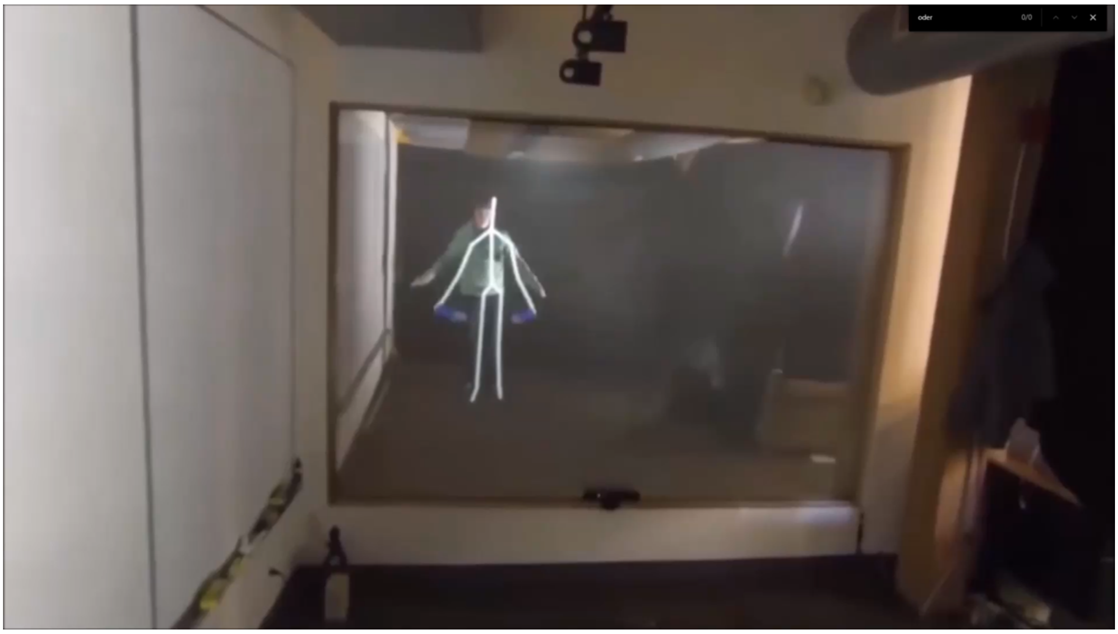


Figure 3.2: Implemented visual perspectives by Anderson et al. [1] for YouMove.

Figure 3.3: VR Dance Trainer by Chan et al. [2]

After each stage, feedback is given. This view provides multiple views for the four keyframes: a stick figure of the teacher and student superimposed, a video of the teacher's demonstration and a video of the moves of the student. Thus, YouMove provides only the exo-centric visual perspective on the guidance visualisation.

VR Dance Trainer. Chan et al. [2] facilitate a 3D screen for rendering the avatars. The student is tracked by an optical motion capturing system. In the first phase (demonstration) only the teacher is shown on the screen. In the second phase (practise) teacher and student can be seen simultaneously, standing side by side. Additionally, both avatars are mirrored, so the exo-centric viewer can observe the avatars from the back and the front at the same time, compare figure 3.3. In the final phase (feedback) the student sees the teacher and the students performance in slow motion side by side.

Onebody by Hoang et al. [10] show the guidance visualisation in an ego-centric perspective, with an HMD. The teacher is projected inside the body of the student, compare figure 3.4 and figure 3.10. Both teacher and student are tracked by skeletal tracking, using a Microsoft Kinect. The visualisation of the teacher is attached to the hip of the student. To overcome different body sizes, the avatars are normalised and scaled to the size of the student. For a second condition, an exo-centric view is provided with an HMD, too. There the teacher stands in front

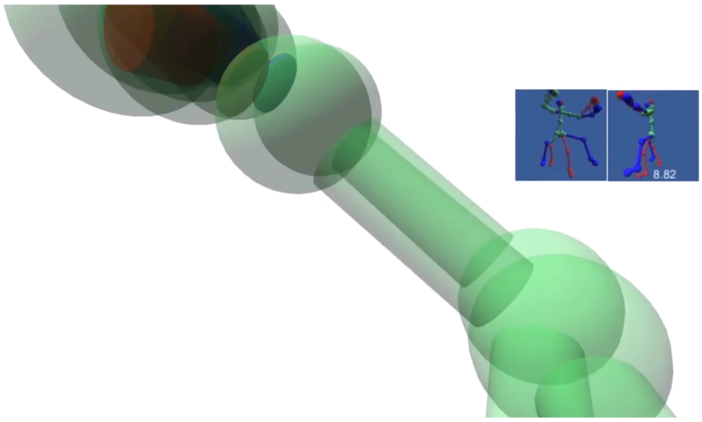


Figure 3.4: Ego-centric perspective on the right arm in Onebody [10]. Teachers and students arm can be seen in overlay [10].

Figure 3.5: Ego-centric guidance visualisation on the hand of the student by [22] in LightGuide.

of the student. **LightGuide** [22] achieved an ego-centric guidance without using a HMD. The student stands under a depth camera and a projector. The depth camera tracks the hand of the student. This position is utilised and the guidance visualisation is projected directly on the user's hand, compare figure 3.5. During the evaluation of the system, four conditions used this ego-centric position. The remaining conditions are shown exo-centric on a screen or directly on the hand.
Physio@Home by Tang et al. [24] used a Microsoft Kinect in combination with

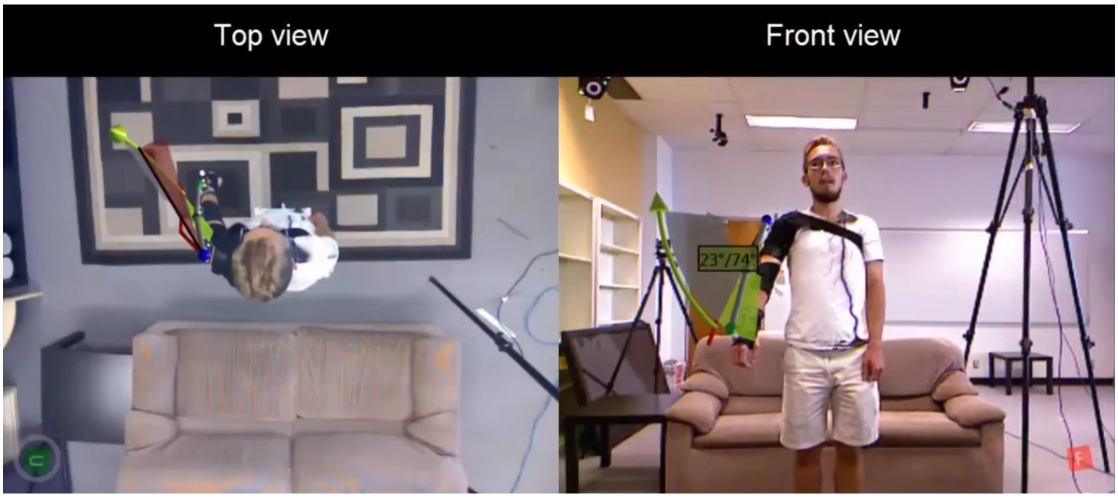
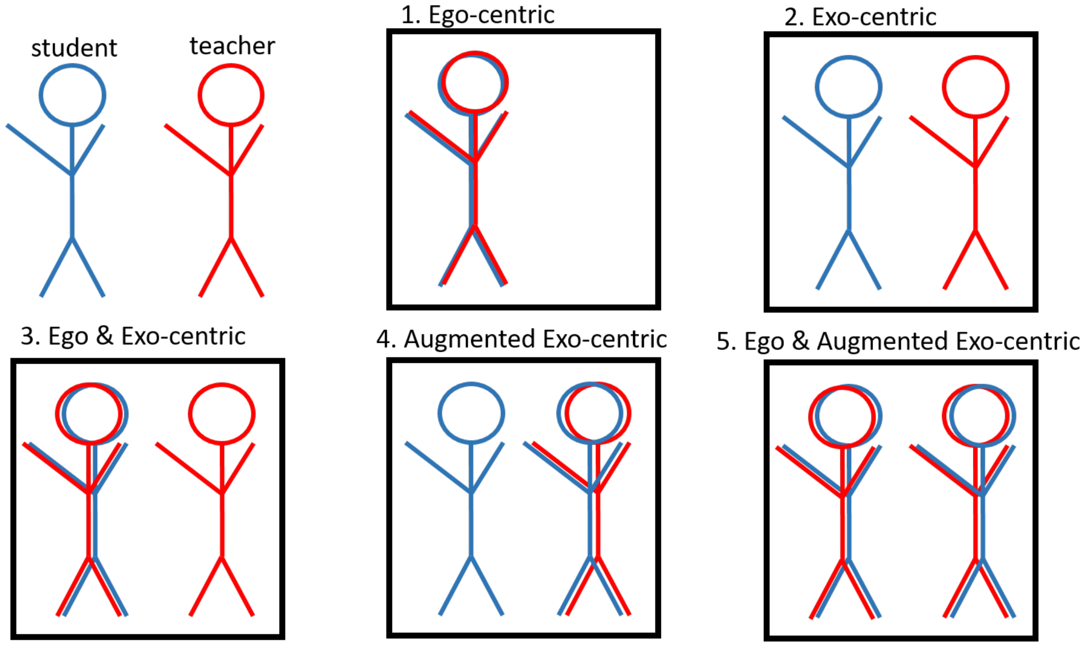


Figure 3.6: Top and front view of Physio@Home [24].

a screen forming together an augmented reality mirror. The participant stands in front of the screen and is tracked by the Kinect. On the screen, the participant is augmented by moving instructions. The visual perspective on the guidance visualisation is exo-centric. Tang et al. faced the problem, that the front-facing projection lacks 3D queues. Their solution was multi-camera views. A second view is provided from above, aiming to help the participant to maintain a correct angle of the arm. In the study, four conditions were examined: videoSingle, videoMulti, wedgeSingle and wedgeMulti. Here, video and wedge indicate the guidance visualisation, single and multi indicate the numbers of visualisations.

Conclusion



	Ego-centric guidance	Exo-centric guidance	Ego-centric self representation	Exo-centric self representation
Perspective (1)	X		X	
Perspective (2)		X	X	
Perspective (3)	X	X	X	
Perspective (4)		X	X	X
(Perspective (5))	X	X	X	X

Figure 3.7: Top: all possible perspective given one student and one teacher with a maximum of two representations. Bottom: detailed differences between the conditions.

Given one student and one teacher there are five possible visual perspectives, compare figure 3.7 1-5. Perspective (1) is the ego-centric view like seen in LightGuide [22], Onebody [10] and in the Tai Chi Trainer [17]. Perspective (2) is the exo-centric visual perspective which in all systems were present. Perspective (3) is a combination of (1) and (2) where the student sees the teacher superimposed on the own body and additionally in front. In perspective (4) the student sees the teacher superimposed by the own body but is not superimposed himself. Finally, the combination of (3) and (4) is shown in perspective (5). Chua et al. used (5) in condition (d) and (e) in their Tai Chi Training system.

In general, all five visual perspectives are interesting to investigate. But (5) promises to overwhelm the student with too much information. For this reason, visual perspective 1-4 will be conditions in the proposed study design.

Another important lesson learned refers to the *mirror effect*. A student standing in front of the teacher needs to mirror the movement before the student can perform the movement. Means, the teacher rises e.g. the left arm, the student wants to lift the right arm instinctively. Furthermore, given a fixed viewpoint on the teacher can lead to depth perception misinterpretations. E.g., the teacher raises the left arm not straight but a little bit to the side, it is hard to grasp the correct angle of the rising of the arm. Chua et al. overcome this with multiple representations of the teacher surrounding the student, though the student can look on the teacher from different directions, Tang et al. provided a front and top view of the guidance visualisation and the Chan et al. show front and back representations of the teacher. For Hoang et al. and Sodhi et al. this was not a problem for their ego-centric perspective. This made clear that in exo-centric visual perspectives, multiple views on the guidance visualisation must be provided.

To overcome the mirror issue, in the proposed study design, the student can move freely around the guidance visualisation to observe it from multiple viewpoints.

3.2 Study Task

	Tai Chi Trainer	YouMove	VR Dance Trainer	OneBody	LightGuide	Physio@Home
Task	Tai Chi	Dance (Ballet), abstract	Dance (HipHop)	Martial Arts	Abstract	Shoulder rehab

Table 3.4: Overview: tasks used by the systems.

The task in Chua's [17] **Tai Chi trainer** is – as the name indicates – a Tai Chi motion. A professional Tai Chi trainer was invited to perform a so-called *Tai Chi form* and recorded offline. This *form* was segmented in four ca. 20 seconds long sequences (Motion 1-4). "Motion 1 featured simple hand movements and a 90 degree turn to the right; motion 2 had little hand motion, a 180-degree turn of the feet, and some turning of the upper body; motion 3 featured slow hand motions, large movement of the feet, and some turning of the thorax; and motion 4 featured swift hand movement and turning of the thorax and hips but little movement of the feet." [17]. Error measurements indicated that all motion but motion 1 had the same difficulty, being significantly easier. This movements can be classified as *sequential movements* according to chapter 2 equation 2.2. 40 volunteers conducted the movements in a study to evaluate their system. They randomized the condition (compare figure 3.1 condition a-e) and the motion to minimize learning effects. For each motion and condition pair, the Tai Chi student were asked to match the Tai Chi teachers demonstration during twelve repetitions.

YouMove by Anderson et al. [1] is a movement training system, suitable for a vast range of moves. A movement can be recorded and then edited by an authoring tool. After the editing, the movement is added to the internal library, from which it can be chosen by a student. For the study itself, an author or the paper – though no professional – recorded four movements. Two of them from ballet and two abstract movements. The authors decided to variate the difficulty of the task, namely "the ballet movements [...] were easier to conceptualise and required only moderate movement.", while "the abstract movements were more difficult to perform, as they were a series of postures with no clear structure and required substantial movement." The movements consisted of four keyframes. Keyframes are important points during the movements, determined and set by the person recording a movement.

Jacky Chan et al. developed a **VR Dance trainer** [2]. For their system, they invited professional Hip-Hop dancers and recorded their movements. The learner of the movement can choose a movement out of a database. One movement lasts around 2 seconds. First, the teacher appears and the student watches the demonstration, while he can adjust the demonstration speed and viewpoint. After that, the student practices the dance moves by mimicking the teacher's avatar. Finally, the student can see a slow-motion replay of the performance. One hole session takes 15 minutes.

Onebody by Hoang et al. [10] is designed for sports or physical activity training like yoga, dance or martial arts. For the study, they used a Martial Arts movement. Each participant started with a training session in which a remote instructor teaches a posture physically and verbally. Verbal feedback was given and the training repeated until the student was confident. After that, the final

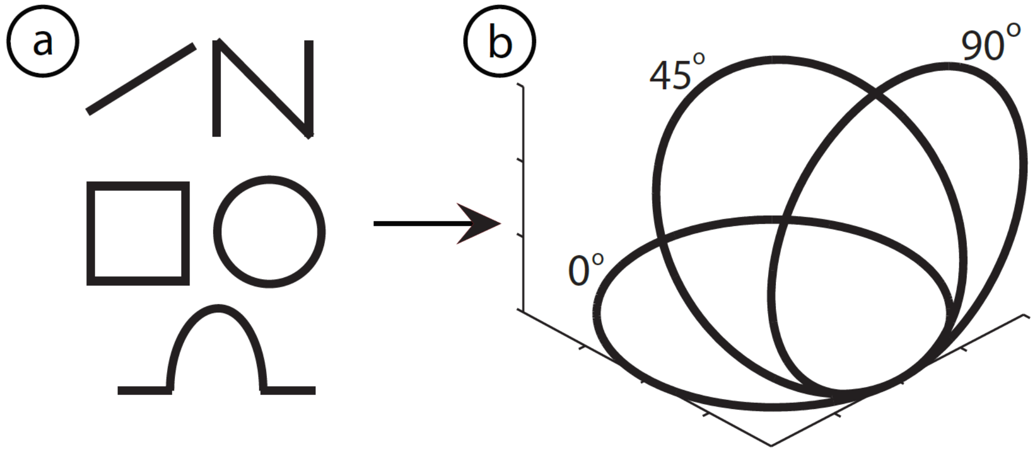


Figure 3.8: Movements schema used by [22] in LightGuide.

posture was recorded. This applied for four postures with different complexities.

In contrast to all other systems, **Lightguide** by Sodhi et al. [22] build purely on abstract movements to evaluate their system. Sodhi et al. focus on single-arm movements, directed by a guidance visualisation on the hand. The participants perform five movements, namely a line, an "N", a square, a circle and a bowed line, compare figure 3.8 left. All movements were performed in three different angles like shown in figure 3.8 right.
These movements were performed with six conditions resulting in 90 dataset per participant (6 conditions x 5 path x 3 angles).
Physio@Home by Tang et al. [24] aims to support rehab exercises for patients at home. Tang et al. chose the shoulder segment as study object, because participants could do this movement easily while standing, and the ball-socket joint gives more DOFs than e.g. the knee. In the study itself, 4 tasks had to be completed: straight, angled, elbow and combo.

straight "Abduction of arm along the frontal plane up to shoulder level, followed by adduction of arm back to the participant's side. This is a simple frontal plane exercise. Angled."

angled "Abduction of the arm at 45° from the frontal plane, followed by adduction back to the side. This is an angled variation of the Straight exercise, where interpreting the angle may be difficult."

elbow "External rotation of forearm away from the center of the participant's body until 90° from the sagittal plane, followed by an internal rotation back to center. This exercise requires the participant to keep their elbow tucked against their side and is a difficult exercise to understand without depth cues (i.e., with just a frontal view)."

combo "Abduction of the arm along the frontal plane up to shoulder level, internal rotation of the arm until pointing forward, followed by an external rotation of the arm back to the frontal plane, and adduction of the arm back to the participant's side. This is a more complex exercise than the previous three, involving many components."

Conclusion

As we see in this section, there is a verity of tasks to evaluate motor learning. Namely, a Tai Chi form (Tai Chi Trainer [17]), dance movements (YouMove [1] and VR Dance Trainer [2]), physiological rehab movements (Physio@home [24]), abstract movements (YouMove, LightGuide [22]) and Martial Arts (Onebody [10]). For the latter is to mention that the evaluation aimed for postures and not movements. To gain valid data some choose to have tasks with comparable complexity ([17], [1] inside Ballet tasks and inside abstract tasks, [22]), therefore e.g. one Tai Chi form was taken and split into four sub forms. Physio@home and YouMove (between ballet and abstract) chose to have different complexities of tasks. All but Tai Chi Trainer proofed to be valid for evaluation of movements. But the authors of Tai Chi Trainer see the reason in the hardware performance, which could be overcome with today's technology. All systems but LightGuide aim to be a teaching system for real-world tasks. This thesis aims to evaluate the influence of the perspective on guidance visualisations itself, therefore, both tasks – abstract and real-world tasks – seem suitable for the proposed study design.

Chua et al. included four visual perspectives out of the five perspectives taken into account for the proposed study design in chapter 4. This invites to reuse their task. Additionally this task has a high ecological validity. Tai Chi is also a well established task in studies for motor learning [1, 9, 13] Furthermore, for comparing the student's performance of the movement, it es inevitable to use movements with the same complexity. Chua et al. solved this by splitting on Tai Chi form into four subforms resulting in movements with nearly the same complexity. For this reason, the task in the proposed study design will be a split Tai Chi form.

3.3 Guidance Visualisation

	Tai Chi Trainer	YouMove	VR Dance Trainer	OneBody	LightGuide	Physio@Home
Guidance Visualisation	hr avatar, wireframe, mimic avatar	Stick figure, mimic avatar	hr figure, mimic avatar	Stick figure, mimic avatar	Indicators, follow/mimic	Indicators

Table 3.5: Overview: guidance visualisations used by the systems.

Chua et al.'s **Tai Chi training** scenario take place in pure VR using an HMD. The lessons are conducted in a virtual pergola standing in nature or park environment. In condition a to c (compare figure 3.1), the student and the teacher are rendered as non-transparent, high realism degree avatars. However, in condition d and e, the visual appearance changes. In condition (d) the student is still rendered normally but the teacher is now represented as a red wireframe. Condition (e) was introduced late in the development of the system based on early subject feedback. The students here are rendered as green wireframe avatars and the teacher as red wireframe avatars.
In terms of guiding technique, the student is presented the pre-recorded motion of the teacher. The teacher avatar performs the motion in question and the student tries to mimic this motion. During the mimicking process, no feedback is provided to the student. But in condition (d) and (e) the teacher is superimposed on the student's avatar. This can be interpreted as feedback since the difference between the students e.g. arm position and the teacher's arm position can be seen easily.

The visualisation of the teacher in **YouMove** [1] is rendered in a low degree of realism – only a stick figure in yellow. During the feedback screen after each stage, the student is also rendered as a stick figure but in blue. For teaching, YouMove utilises the above mentioned five training stages. In the first stage (Demonstration), a video of the teacher performing the movements is shown. The student only sees the video and not the reflection of their own body. The second stage (Posture guide) a video and the stick figure of the student are shown, but still not the own reflection. The student is asked to match they postures at specific keyframes, where the demonstration stops. After matching the posture of a keyframe, the demonstration moves on to the next keyframe. In the third stage (Movement Guide) the demonstration no longer stops at the keyframes. Furthermore, the reflection and the projection are visible simultaneously, the student is superimposed by the teachers stick figure. Stage four (Mirror) is like the name indicates: the student stands in front of the mirror seeing only his reflection but additionally

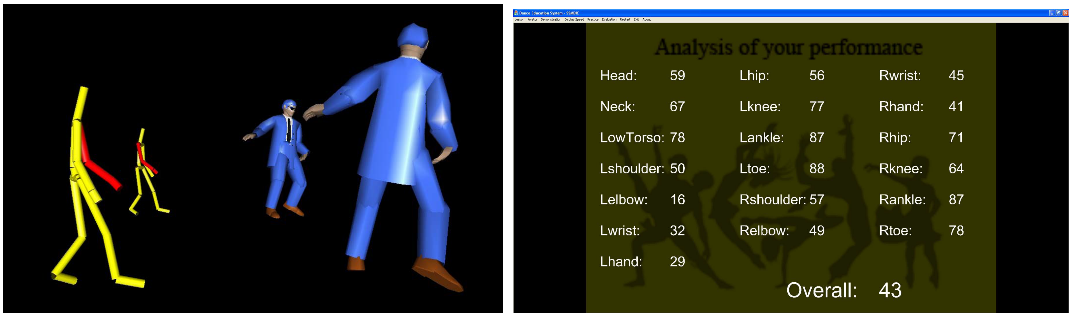


Figure 3.9: Left: student (left) and guidance visualisation (right) by [2] for VR Dance Trainer, right: offline feedback screen ibidem.

audio queues are provided. In the last stage (On your own) even the mirror is removed and the student performs the movements without any guidance.
In terms of feedback, Anderson et al. state "[...] the availability and modality of feedback can greatly impact skill acquisition". To match this, feedback is a part of YouMove. During the second stage (Demonstration), red circles indicate misplaced limbs. The bigger the circle the bigger the displacement. Additionally, a side view on the scene is faded in. At the end of each stage, a so-called "summary feedback" is provided. Here both stick figures of the teacher and student are superimposed and a video of both can be seen. Eventually, a score is provided, which will be explained in the next section.

Chan et al. **VR Dance Trainer** [2] choose two different visual representation. The teacher is shown as a high realism degree avatar, while the student is shown as a low realism degree stick figure. The students stick figure serves additionally as feedback regarding the student's performance. The body parts are coloured from green (perfect motion match) over yellow (acceptable motion match) to red (poor motion match). In all phases, these visual representations stay consistent. In the first phase (demonstration) no feedback is given, in the second phase (practice) the student gets immediate feedback in the colour code described above. The last phase provides a slow-motion replay showing with feedback in the same colour code. Additionally, the student can get feedback by the scoreboard see figure 3.9 right. Here all body parts are shown with a numeric indicator (0-100) how good the motion was performed on the specific body part.
Onebody [10] use low realism degree avatars. Both, the student and the teacher are visualised by stick figures. The teachers avatar is red, the students avatar is blue. For feedback, if the students joints are matching joints of the teacher, these joints turn from blue to green, like shown in figure 3.10 left. Figure 3.10 right shows the scene from the first person perspective. The guidance itself takes place

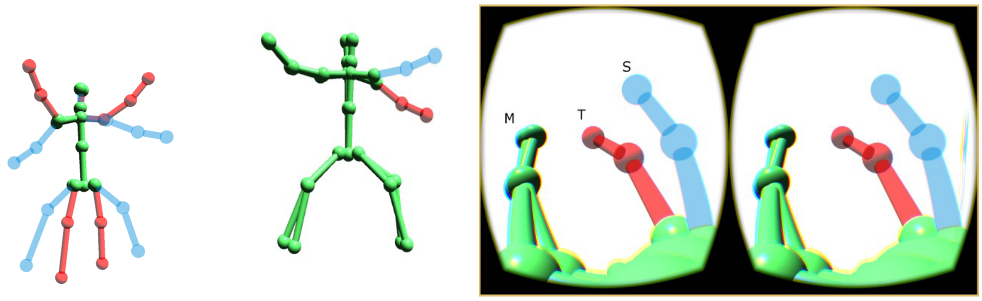


Figure 3.10: Left: student avatar (blue) and teacher avatar (red). Green limbs are matching limbs. Right: students view on the scene. Onebody [10].

	Video	*Skype*	*VR-3PP*	*Onebody*
Synchronous Interaction	No	Yes	Yes	Yes
VR Medium	No	No	Yes	Yes
1st Person View	No	No	No	Yes

Figure 3.11: Training methods and their differences used in the study to evaluate Onebody [10].

in real-time. The remote teacher can give instructions by performing the postures and additionally verbally. The student mimics the postures and can compare his joint positions with the positions of the teacher to correct his joints. Hoang et al. compare the performance of the students achieved in Onebody with the performance in three other systems, namely with traditional video-based learning, video conferencing and a 3rd person perspective in VR. The latter is very similar to Onebody but the teacher stands in front of the student. The system differs in exactly one aspect to each other, see figure 3.11. **LightGuide** [22] compare six guidance visualisations:

- Follow Spot: Moving light spot with elevation information. System sets pace.
- 3D F-Arrow: 3D arrow indicates the direction to move. System sets pace.
- 3D SG-Arrow: 3D arrow indicates the direction to move. User sets pace.
- 3D Pathlet: a line indicates the direction to move next. Current position indicated by a red spot. System sets pace.
- Video on hand: Instruction video projected on the hand. User sets pace.

- Video on screen: Traditional instruction video. User sets pace.

The main idea behind a guidance visualisation directly on the student's body itself is that the student can concentrate on the body part in question and not share them with an instruction medium. The last condition proofed to be better than the second last, supporting this thesis. The student gets instant feedback by comparing the indicators with the hand. A larger offset results in stronger indication in real-time.

Physio@Home by Tang et al. [24] use two guidance techniques. In the first, a video shows the movements instructions, the second they called "the wedge". The latter is a 3D plane showing the way of movement with direction indication. As feedback, the current angle of the arm is shown. On figure 3.6 "the wedge" is depicted.

Conclusion

The visual appearance of the guidance visualisations differ. The Tai Chi Trainer, YouMove, VR Dance Trainer and Onebody use person-shaped avatars. These avatars perform a movement and the student mimics these movements. LightGuide and Physio@Home use direction indicators to visualise the movement in question. Movement indicators like arrows are less obtrusive for single body-part movements but become overwhelming for full-body instructions [22].
This work aims to analyse full-body movements, therefore I choose here as visual appearance a person-shaped avatar. The degree of realism of the avatar rage from stick figures to high realistic-looking persons. Since Weber [27] showed a preference for realistic avatars I hereby choose high realism degree avatars as guidance visualisations.
Besides, there is a difference in completion time based on the guidance technique. If the student sees the movement beforehand completion time is significantly lower. If the student grasps the movement by performing it, completion time is significantly higher. A study must ensure that for all conditions the student could see the movement before performing the movement or not. Comparing with the real world, normally the student sees the movement beforehand. This is why my study will provide pre-watching the demonstration of a movement, too.
Sodhi et al. state, that it is hard to compare self-paced and system paced movements. To ensure comparability, only system paced movements will be evaluated. This ensures also only to evaluate the precision of movements.
In all systems, feedback plays a role. Immediate feedback as well as aftermath feedback (online vs. offline feedback). For comparability, feedback must be con-

sistent over all conditions. This work does claim to provide full training system but an evaluation tool. For this reason, only immediate feedback during learning will be provided.
To ensure, that only the visual perspectives and not feedback are evaluated, the same feedback method will be provided in all conditions. I choose here the overlay feedback without colour coding, where the student align his/her body with the body of the teacher avatar.
Eventually, all authors point out the importance of scaling. To correctly match teacher and student, the teacher must be scaled to the students body.

Weber showed that high realism avatars suit best for the task, therefore the guidance visualisation will be a high realism avatar. For comparability of the study conditions, all tasks will be shown before hand and be paced by the system and not the student. Only immediate feedback will be provided by the overlaying teacher avatar. Finally, to overcome the body size differences, the teacher will be scaled to the student's height.

3.4 Variables

	Tai Chi Trainer	YouMove	VR Dance Trainer	OneBody	LightGuide	Physio@Home
Variables	Perspectives, performance measure	VR/Video, performance	Video/VR, performance	Training method, performance	Visualisations, Perspective, Performance	Visualisation, performance

Table 3.6: Overview: used variables to evaluate the systems.

In the **Tai Chi Trainer** by Chua et al. [17], the student's task is to mimic the motion of a pre-recorded teacher. While the independent variables are the above-mentioned perspectives (conditions a-e) and the motions, the dependent variable is the precision of the performed movements. To measure this precision, twelve bones of the students were tracked, namely: upper and lower arms, hands, upper and lower legs, and feet. The bones are hierarchically structured with a parent and child end. After normalization of the parent's position and bone length, an error was calculated. This error can be seen as the euclidean distance between the teacher's and the student's position of a bone, calculated for every frame of the roughly 20 seconds with a rate of 60 fps. The last four out of twelve trials were considered. This measure is an implementation of chapter 2 equation 2.2. Chua et al. found two major faults in this method of precision determination:

yaw shift and time shift. To overcome the first, the initial position of the student was taken into account, the latter was fixed with a time frame comparison of 120 frames. Eventually, this result was normalized with the difficulty (average number of errors per motion) of the motion task. This will be discussed in detail in the master's projects report. In addition, a post-questionnaire was conducted where the students were asked to rate the difficulty of the representations. The results are discussed in the next section.

While Chua et al. calculate the error of the performed movement over all tracked limbs, Anderson et al. (**YouMove**) take one single joint with the greatest error and – even more constricting – only the keyframe joints (important joints, specified by the teacher) are taken into consideration. The dependent variable is a score between 0 and 10. An offset 15 cm results in a score of 7.5 and no error results in a score of 10. The offset is simply the Euclidean distance. This error measurement corresponds also to chapter 2 equation 2.2. To overcome time shift errors, a window of 0.5 seconds is added. If the teacher specified that timing is important, this window is halved, if precision is set as important, 15cm offset results in a score of 7.5. "[]This values are determined by experimentation".

Chan et al. evaluate the student's performance with the **VR Dance Trainer**. Therefore they specified 19 body parts (compare figure 3.9 right) and calculate a score between 0 and 100. The average over the 19 numeric indicators results in the overall performance. Before and after the training session of one move, this score was calculated. Additionally, a post-course survey asked the student specific questions about the system. This survey had the aim to evaluate if the "[...] system is interesting and able to motivate subjects to learn." [2] and if the "[...] the system can provide them an easy way to learn" ibidem.

Hoang et al. [10] use six different measures to evaluate **Onebody**. *Accuracy*, *completion time*, *instructors score*, *ease to understand*, *perceived precision* and *preference*. For *accuracy* the angles of limbs of the student and teacher were compared. For *completion time* the time between start and "[...] the student feeling confident" was measured, but caps at 2 minutes (compare chapter 2.2). *Instructors score* is a subjective score of the instructor after each posture. The data for the last three measures were gathered by post questionnaires. These scores were calculated for all four independent variables (video, video conference, 3rd person, Onebody) Sodhi et al [22] compare in **LightGuides** five conditions where the instruction are directly on the body of the user with a baseline condition where the instruction is on a screen. To measure the performance of the student, Sodhi et al. developed two measures: movement accuracy and movement times. Movement accuracy describes the absolute euclidean distance from the closest point. Movement times is divided into two sections by the reason that half of the conditions were self-timed and the in other half, the pace was set by the system. For the self-timed conditions, the completion time (compare chapter 2.2) is taken as a measure, for system paced conditions the time before or after is taken as a measure. Though independent variable are the conditions, dependent variables are performance.

Physio@Home by Tang et al. [24] use three performance measures to evaluate the student's performance. Two distance measures, one for the elbow and one for the hand, and one angle measure for a maximum of rotation. They ignored speed as a measure because they were mainly interested in the precision of the movements. Additionally, two subjective measures were gathered: perceived accuracy and preference for method.

Conclusion

All previously mentioned works use a performance measurement. These measurements itself varies but have a precision score in common. Some [1] [2] [24] [22] specify important body parts or give weights on body parts. The error is measured in the euclidean distance. To overcome timing issues, an time frame between teachers and students movement apply. Besides, a second measurement is used: time. Completion time measures the time a student needs to feel confident to perform the movement. As subjective measurements, perceived accuracy and preference of method is widely used. Additionally, Onebody asks the participants to rate the ease to understand.

> For the proposed study design a distance-based performance measure with a 0.5s time shift will be used which proofed to be suitable in [1]. Furthermore, a completion time could give objective insights on perceived precision and speed of learning. As subjective scores, ease to understand, preferred method, perceived precision will be taken.

3.5 Results

	Tai Chi Trainer	YouMove	VR Dance Trainer	OneBody	LightGuide	Physio@Home
Results	No difference in performance	VR better than video	VR better than video	Ego better than exo	Ego better than exo	Multi view better than single view

Table 3.7: Overview: results of related work.

Tai Chi Trainer. Chua et al. compare different perspectives (conditions a-e) on the teacher and the student. The independent variables are the above-mentioned precision of the student's performance. By comparing the results they found condition (a) (One on One), (b) (Four Teachers), (c) (Side by Side) and (e) (Superimposition 2) aim the same precision. Only (e) (Superimposition 1) aimed significantly worse than the others. At the same time, the questionnaire indicated that the subjective difficulty of condition (d) and (e) (Superimposition 1 and 2) was the highest. "In fact, all of the subjects who tried Superimposition 2 thought it was the most difficult. Interestingly, although subjects considered Superimposition 2 very difficult compared to the other layouts, the average error on that layout was not significantly greater than the other non-superimposed layouts." [17]. The Authors argue this result as follows:

- simultaneously watching the teacher and performing own movements could interfere with each other.
  Chua et al. suggest choosing wisely for the task to suit into VR training.
- latency and performance correlate strongly. A lower latency could lead to better performance.
  Since the Tai Chi Trainer was developed in 2003, there is a large improvement in latency nowadays.
- To reduce latency a low polygon count on the high realism degree avatars was used. More polygons could lead to better performance.
  The system was run on a Pentium 3 processor. Today's graphics cards and processors are way above this mark, a higher polygon count is easily achievable.
- The field of view was very small.
  Today's VR HMDs provide a higher field of view (e.g. HTC Vive, 110°).

Anderson et al. compared **YouMove** with traditional video training, resulting in the independent variables *YouMove* and *Video*. The study was conducted with eight participants in a two-factor repeated-measures design. Each participant had one ballet and one abstract task with both conditions. *YouMove* scored significantly better than *Video* by a factor of 2.

Chan et al. [2] investigated three topics with the **VR Dance Trainer**. First the learning outcome, to proof if the student got better with the system. Second, the "Arousing Interest" to investigate whether the system motivates the students to learn. Eventually, they compared the system with a traditional self-learning method. Comparing the baseline score before training with the VR Dance Trainer with the score after the training session proofed a significantly better performance. The post-survey is interpreted by the authors with "Overall speaking, the subjects enjoy learning dance with our proposed system." ibidem. To compare the system with a traditional learning method, a control group conducted the same study only with the demonstration and no feedback. The baseline scores showed no significant difference between the two groups but the post-training scores did. Questionable remains if a recording of a professional dancer rendered as a high realism degree avatar on a 3D screen can be called a traditional dance learning method.

Onebody [10] proved to be significantly better over the other training measures in terms of *accuracy*. Interestingly, no significant difference was found between the exo-centric 3rd person view and video conference. On the same time, Onebody has a higher *completion time* than the other systems. The instructor's score showed no significance between the methods. Valuable for this work is, that the ego-centric perspective seems to be slightly better to *understand* than the exo-centric perspective, but not significant. Hence, both ego-centric and exo-centric perspectives are significantly harder to understand that the video-based methods. Furthermore, the *perceived precision* ego-centric perspective is significantly higher than the exo-centric perspective, but nearly on the same level as the video conference method. Eventually, Onebody is more *preferred* by the participants of the study than the 3rd person view. Hoang et al. conclude "[...] that synchronous training and 1st person view has a positive effect on posture accuracy." [10] If this applies also to movements could be interesting to investigate.

The main finding of Sodhi et al. while evaluating **LightGuide** is an 85% higher accuracy of ego-centric on body guidance visualisations compared to exo-centric video instructions. Per student 90 datasets are generated (6 conditions x 5 path x 3 angles). The performance measure they applied, the conditions scored (best to worst): Follow Spot < 3D F-Arrow < 3d G-Arrow < 3D Pathlet < Video Hand < Video Screen. Especially the relation between video on screen and video hand shows the importance of attention during guidance instruction. The instruction on the body-part it self scored better than seeing the instruction on a screen. In terms of *movement times*, both video conditions lead to lower movement times. Sodhi et al. see the reason therefore, that it makes a difference if the student sees the whole path beforehand than or figuring out the movement as they moved along. Additionally, in an interview, participants of the study stated that self-paced guidance are is subjectively preferred over system paced. For system paced movement speed, 30mm per second scored best in a pilot test. Further, they suggest, to plan regular recovery rests to exclude fatigue effects.

Physio@home by Tang et al. [24] investigated two subjective measures. The perceived accuracy ranked videoSingle < videoMulti < WedgeSingle < wedgeMulti. For the preferred method aimed for no clear preference. The study results show a higher performance with the multi-view perspectives and also the wedge visualisation. The wedge visualisation in combination with multi-view perspective achieved the highest performance. The authors see the reason in the ability to grasp the correct angle is much easier with the multi-view perspective. Furthermore, the "corrective" feedback of the wedge encouraged the participants to correct themself. Tang et al. further state, that visualisations should contain as least as necessary information to not overwhelm the practitioners.

Conclusion

LightGuide showed that ego-centric guidance performed better than exo-centric. On reason, they see in the shift of attention. When the student can focus on the body part which is to move itself, the non-shared attention plays a positive role in learning outcome. The ego-centric Onebody system proofed to be better than the exo-centric perspectives. Since Onebody only evaluates posture guidance, it could be interesting to find out if this also applies for movement guidance. Comparing all systems that implemented ego-centric and exo-centric perspectives concludes as following: Onebody and LightGuide achieved better results with the ego-centric view. The Tai Chi Trainer saw no difference between the two perspectives but see the problem in the hardware available in 2003. Covaci et al. (Free Throw Simulator) [6] and Kojima et al. (Training Physical Skill) [12] implemented both visual perspectives in but could not find a clear preference for one over the other. Hachimura et al. (MR Dance Trainer) [7] and Sousa et al. (SleeveAR) [23] did not evaluate the difference between the two visual perspectives. This makes clear, that further investigations on visual perspectives are necessary. This work aims to close this research gap.

3.6 Summary

Chapter 2 defined the scope for this work. With this scope in mind, this chapter defined – based on related work – the parameters for study design. The parameters are clustered in four main topics: visual perspective, task, guidance visualisation and measures. In terms of guidance visualisation, we saw five possible visual perspectives given one teacher and one student. Four out of five will serve as independent variables for the study. Furthermore, it becomes clear that to overcome the mirror effect, the student must be able to move freely around the guidance visualisation. The task will be adopted from Chua et al.'s Tai Chi Trainer. For the guidance visualisation, a scaled high realism degree avatar will be used and the system paced movement is shown beforehand. Feedback will not be part of the system, but overlying the student's own body with the body of the teacher can be seen as a feedback method. For objective evaluation, euclidean distance-based error measurements and time measurements will be applied. As subjective scores, the ease to understand, perceived precision and preferred method will be gathered. Finally, the results of the discussed related work showed clearly the need for further investigations on the topic of visual perspectives and motor learning.

4 Proposed Study Design

Given the scope from chapter 2 and the parameters from chapter 3, this chapter proposes a study design. The study aims to produce data to answer the main research question from chapter 1:

MRQ Does the visual perspective on a virtual guidance visualisation have an influence on motor learning in MR environments.

4.1 Setup

For the study, movement training system will be implemented. This movement training system will include a virtual reality HMD and motion-tracking technology. The student will be tracked with this motion capturing technology and with the resulting information, the student's avatar will be rendered as a high realism degree avatar. Likewise, the teacher avatar will be rendered, but not on the base of live motion tracking data. A professional Tai Chi trainer will be invited and a Tai Chi form will be recorded. This form will be split into 4 sub-forms. Furthermore, the teacher avatar will be scaled to the size of the student's avatar. To overcome the mirror issue, the student is allowed to move freely around the teachers' avatar. During the students' performance, the movement will be recorded and analysed in the aftermath with the discussed performance measure.

4.2 Procedure

The study is conducted in a within-subject design with counterbalancing, which results in 32 participants, compare table 4.1. The student starts with the first visual perspective. The teacher appears and performs the movement, while the student is watching the performance. After the first demonstration, the student can train the motion simultaneously until the student is feeling confident, but caps after an amount of time which a pilot study has to reveal. After one of these cases, the final movement will be recorded, the student gets a short recovery rest. Then the next visual perspective starts. This process continues throughout all

Ego-centric	Exo-centric	Augmented exo-centric	Ego & Exo centric
Tai Chi sub-from 1	Tai Chi sub-from 2	Tai Chi sub-from 3	Tai Chi sub-from 4

Table 4.1: Study conduction schema.

four conditions. Subsequent to the actual study, the post questionnaires will be answered.

4.3 Outlook

The next step is to implement the study setup, but still, there are some parameters to refine on. This takes place in the master's project. In this part, the technology and algorithms will be in the main focus. For technology, VR HMD will be compared and a decision will be made. Similar to motion tracking technologies. After this, algorithms for comparing two movements will be evaluated. When decisions are made, the implementation of the system will start. Eventually, a pilot testing session will take place followed by further refinements. The milestones for the master's project are:

- Hardware requirements: 17.1.2020
- Software Requirements: 24.1.2020
- Implementation Start: 27.1.2020
- Pilot Study: 2.3.2020
- Refinements Implementation: 13.3.2020
- Master's Project Presentation and Report: 23.3.2020

The thesis will follow in the summer semester 2020. The results are planned to be published at a conference in 2020, compare figure 4.1.
For the master's thesis, a study will be conducted. The study will generate data suitable to answer the research question. This data will be analysed in detail and eventually be used to answer the research question.

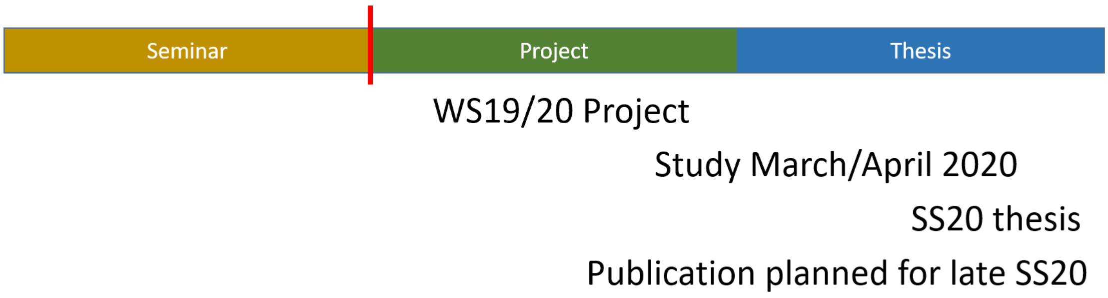


Figure 4.1: Timetable for the master's thesis.

Bibliography

[1] Fraser Anderson, Tovi Grossman, Justin Matejka, and George Fitzmaurice. YouMove: Enhancing Movement Training with an Augmented Reality Mirror. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*, UIST '13, pages 311–320, New York, NY, USA, 2013. ACM.

[2] J C P Chan, H Leung, J K T Tang, and T Komura. A Virtual Reality Dance Training System Using Motion Capture Technology. *IEEE Transactions on Learning Technologies*, 4(2):187–195, apr 2010.

[3] Jacky Chan, Howard Leung, Kai Tai Tang, and Taku Komura. Immersive Performance Training Tools Using Motion Capture Technology. In *Proceedings of the First International Conference on Immersive Telecommunications*, ImmersCom '07, pages 7:1—-7:6, ICST, Brussels, Belgium, Belgium, 2007. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

[4] Winyu Chinthammit, Troy Merritt, Pedersen Scott, Williams Andrew, Visentin Denis, Robert Rowe, and Furness Thomas. Ghostman: Augmented Reality Application for Telerehabilitation and Remote Instruction of a Novel Motor Skill. *BioMed Research International*, 2014, 2014.

[5] Worawat Choensawat, Minako Nakamura, and Kozaburo Hachimura. GenLaban: A Tool for Generating Labanotation from Motion Capture Data. *Multimedia Tools Appl.*, 74(23):10823–10846, dec 2015.

[6] Alexandra Covaci, Anne-Hélène Olivier, and Franck Multon. Third Person View and Guidance for More Natural Motor Behaviour in Immersive Basketball Playing. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*, VRST '14, pages 55–64, New York, NY, USA, 2014. ACM.

[7] K Hachimura, H Kato, and H Tamura. A prototype dance training support system with motion capture and mixed reality technologies. In *RO-MAN*

2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No.04TH8759), pages 217–222, 2004.

[8] Ping-Hsuan Han, Kuan-Wen Chen, Chen-Hsin Hsieh, Yu-Jie Huang, and Yi-Ping Hung. AR-Arm: Augmented Visualization for Guiding Arm Movement in the First-Person Perspective. In *Proceedings of the 7th Augmented Human International Conference 2016*, AH '16, pages 31:1—-31:4, New York, NY, USA, 2016. ACM.

[9] Ping-Hsuan Han, Yang-Sheng Chen, Yilun Zhong, Han-Lei Wang, and Yi-Ping Hung. My Tai-Chi Coaches: An Augmented-learning Tool for Practicing Tai-Chi Chuan. In *Proceedings of the 8th Augmented Human International Conference*, AH '17, pages 25:1—-25:4, New York, NY, USA, 2017. ACM.

[10] Thuong N Hoang, Martin Reinoso, Frank Vetere, and Egemen Tanin. Onebody: Remote Posture Guidance System Using First Person View in Virtual Environment. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*, NordiCHI '16, pages 25:1—-25:10, New York, NY, USA, 2016. ACM.

[11] Nicholas Katzakis, Jonathan Tong, Oscar Ariza, Lihan Chen, Gudrun Klinker, Brigitte Röder, and Frank Steinicke. Stylo and Handifact: Modulating Haptic Perception Through Visualizations for Posture Training in Augmented Reality. In *Proceedings of the 5th Symposium on Spatial User Interaction*, SUI '17, pages 58–67, New York, NY, USA, 2017. ACM.

[12] Taihei Kojima, Atsushi Hiyama, Takahiro Miura, and Michitaka Hirose. Training Archived Physical Skill through Immersive Virtual Environment. In Sakae Yamamoto, editor, *Human Interface and the Management of Information. Information and Knowledge in Applications and Services*, pages 51–58, Cham, 2014. Springer International Publishing.

[13] Taku Komura, Beta Lam, R W H Lau, and Howard Leung. e-Learning Martial Arts. *Lncs*, 4181:239–248, 2006.

[14] Joseph J LaViola. 3D user interfaces: theory and practice , 2017.

[15] J Lieberman and C Breazeal. TIKL: Development of a Wearable Vibrotactile Feedback Suit for Improved Human Motor Learning. *IEEE Transactions on Robotics*, 23(5):919–926, oct 2007.

[16] Paul Milgram and Fumio Kishino. A Taxonomy of Mixed Reality Visual Displays. *IEICE Trans. Information Systems*, vol. E77-D:1321–1329, 1994.

[17] Philo Tan Chua, R Crivella, B Daly, Ning Hu, R Schaaf, D Ventura, T Camill, J Hodgins, and R Pausch. Training for physical tasks in virtual environments: Tai Chi. In *IEEE Virtual Reality, 2003. Proceedings.*, pages 87–94, mar 2003.

[18] Otniel Portillo-Rodriguez, Oscar O Sandoval-Gonzalez, Emanuele Ruffaldi, Rosario Leonardi, Carlo Alberto Avizzano, and Massimo Bergamasco. Real-Time Gesture Recognition, Evaluation and Feed-Forward Correction of a Multimodal Tai-Chi Platform. In Antti Pirhonen and Stephen Brewster, editors, *Haptic and Audio Interaction Design*, pages 30–39, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.

[19] Vijay Rajanna, Patrick Vo, Jerry Barth, Matthew Mjelde, Trevor Grey, Cassandra Oduola, and Tracy Hammond. KinoHaptics: An Automated, Wearable, Haptic Assisted, Physio-therapeutic System for Post-surgery Rehabilitation and Self-care. *Journal of Medical Systems*, 40(3):60, dec 2015.

[20] Stjepan Rajko and Gang Qian. A Hybrid HMM/DPA Adaptive Gesture Recognition Method. In George Bebis, Richard Boyle, Darko Koracin, and Bahram Parvin, editors, *Advances in Visual Computing*, pages 227–234, Berlin, Heidelberg, 2005. Springer Berlin Heidelberg.

[21] Richard A. Schmidt and Tim Lee. *Motor Control and Learning.* Human Kinetics, 5th editio edition, 2011.

[22] Rajinder Sodhi, Hrvoje Benko, and Andrew Wilson. LightGuide: Projected Visualizations for Hand Movement Guidance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 179–188, New York, NY, USA, 2012. ACM.

[23] Maur\'\icio Sousa, João Vieira, Daniel Medeiros, Artur Arsenio, and Joaquim Jorge. SleeveAR: Augmented Reality for Rehabilitation Using Realtime Feedback. In *Proceedings of the 21st International Conference on Intelligent User Interfaces*, IUI '16, pages 175–185, New York, NY, USA, 2016. ACM.

[24] Richard Tang, Xing-Dong Yang, Scott Bateman, Joaquim Jorge, and Anthony Tang. Physio@Home: Exploring Visual Guidance and Feedback Techniques for Physiotherapy Exercises. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 4123–4132, New York, NY, USA, 2015. ACM.

[25] Eduardo Velloso, Andreas Bulling, and Hans Gellersen. MotionMA: Motion Modelling and Analysis by Demonstration. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 1309–1318, New York, NY, USA, 2013. ACM.

[26] Wenbi Wang and Paul Milgram. Dynamic Viewpoint Tethering for Navigation in Large-scale Virtual Environments. *Human Factors*, pages 1862–1866, 2001.

[27] Rebecca Weber. Egocentric Motion Guidance in Mixed Reality - A comparison of visualizations guiding bilateral arm movements, 2018.

[28] Shuo Yan, Gangyi Ding, Zheng Guan, Ningxiao Sun, Hongsong Li, and Longfei Zhang. OutsideMe: Augmenting Dancer's External Self-Image by Using A Mixed Reality System. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '15, pages 965–970, New York, NY, USA, 2015. ACM.

[29] Ungyeon Yang and Gerard Jounghyn Kim. Implementation and Evaluation of "Just Follow Me": An Immersive, VR-Based, Motion-Training System. *Presence*, 11(No. 3):304–323, 2002.

[30] Mitsu Yoshimura, Hideki Murasato, Tamiko Kai, Akira Kuromiya, Kiyoko Yokoyama, and Kozaburo Hachimura. Analysis of Japanese dance movements using motion capture system. *Systems and Computers in Japan*, 37(1):71–82, 2006.