Ego OR Exo:
Comparing Visual Perspectives on Guidance Visualisations for Motor Learning

Masterarbeit

vorgelegt von

Stefan Paul Feyer

an der

Universität Konstanz

Sektion Mathematik und Naturwissenschaft

Fachbereich Informatik und Informationswissenschaft

1. Gutachter: Prof. Dr. Harald Reiterer
2. Gutachter: Dr. Karsten Klein

Konstanz, 2021

Acknowledgements

Thank you, Maximilian Dürr, for your phenomenal unremitting support and the patience of a saint to answer my questions.

Thank you, Prof. Dr. Harald Reiterer, for valuable hints which put me back on track, throughout all my study years. Thank you, for the educational time at the Human-Computer Interaction Group.

Thank you, Gerhard Feyer, for your exceptional sempiternal support, in every aspect.

Thank you, Betzi, being a loyal helper during the development of E(x|g)o, in times where nobody else was allowed to help.

Abstract

Motor learning is a substantial part of life, like in sports, arts or the ergonomic handling of physical load. Motor learning is traditionally performed with the help of a teacher. If a teacher is not available, a digital guidance visualisation in Virtual Reality can be consulted. When motor learning is done together with a human teacher, a learner can watch the teacher's movements from an exo-centric visual perspective (third-person perspective). In contrast, in Virtual Reality, a guidance visualisation can be seen from the ego-centric visual perspective (first-person perspective), too. The change of the visual perspective on the guidance visualisation influences motor learning. However, the empirical evidence about how the change of the visual perspective influences motor learning is low, especially for full-body movements, for tasks which include the ergonomic handling of a physical load and visual perspectives that utilise ego-centric and exo-centric guidance visualisations simultaneously. Furthermore, the field of ego-centric guidance of locomotion movements is unexplored. This master's thesis proposes an experiment to close this research gap. The experiment compares an ego-centric visual perspective, an exo-centric visual perspective and a visual perspective which combines both. The experiment utilises a task which consits of elemental tasks of handling physical load. For the evaluation of the experiment, accuracy measurements, ergonomic measurements, the learner's visual focus and qualitative data is taken. The experiment was evaluated with a pilot study and proved to be suitable to generate the data to close the above-mentioned research gap. First data indicates that the presence of an ego-centric guidance visualisation positively influences accuracy and attracts the visual focus of the learner.

Contents

List of Figures

List of Tables

1. Introduction

The acquisition of movements is a crucial part of human development[1]. Learning movements empowers to be more efficient, faster or more exact (ibid.) for tasks like sports, arts or the ergonomic handling of physical load. The process of learning movements is called motor learning.

Movements can be learned by voyeurism and mimicking: watching and trying it out by oneself. Mastering a movement mostly includes an experienced teacher. A teacher is hardly replaceable because of immediate visual, audible and haptic feedback on a movement performed. However, if a teacher is not available, for example, based on the location or economic reasons, other sources can be used to learn movements. For example, YouTube[1] and other video platforms have become a great source for learning videos for a wide range of purposes. The downside of video recordings is the two dimensional (2D) experience of a three dimensional (3D) movement. Mixed Reality devices can transport the learning process into the digital world. In contrast to video recordings, Mixed Reality can provide the learning experience in 3D. Furthermore, Mixed Reality can provide feedback on the performed movement and has the ability for interactions with the virtual guidance visualisation. Mixed Reality has already proven to be a suitable environment for motor learning for tasks like dancing [2–6], sports [7, 8], rehabilitation [9–13], arts [14–21] and others [22, 23]. However, this master's thesis will focus on Virtual Reality.

In the real world, where the learner and teacher are real persons, the learner sees the teacher, for example, in front of himself/herself. This perspective is called the exo-centric visual perspective. Nevertheless, if we move from the real world to the virtual world of Virtual Reality, we are no longer restricted to the exo-centric visual perspective. The teacher can be rendered inside the learner's body, allowing the learner to see the teacher from an ego-centric visual perspective. The change from the exo-centric to the ego-centric visual perspective potentially influences motor learning, which is shown by previous research. For example, AR-Arm [14] lets the learner experience the movements from an ego-centric perspective. YouMove [2] teaches dance from an exo-centric perspective. OneBody [20], LightGuide [23], Mixed Reality Dance Trainer [6], Free Throw Simulator [7], Training Physical Skills [8], Sleeve AR [12] and Tai Chi Trainer [21] use both visual perspectives or a combination of them. However, only OneBody, LightGuide and Tai Chi Trainer found differences between the visual perspectives.

Another topic where Virtual Reality could be a valuable helper is the ergonomic conduction of movements while handling a physical load [24, 25]. The handling of physical load is part of working routines and everyday life. Handling physical load in the correct ergonomic conduct in working routines can prevent injuries. However, a kinaesthetics teacher is not always accessible, for example, for economic reasons. The influence of the visual perspective on a virtual guidance visualisation teaching the handling of physical load in Virtual Reality is sparsely investigated. Especially, locomotion movements like walking or carrying in the ego-centric perspective are left out. The lack of research on the influence of

---

[1] https://www.youtube.com/, accessed 17.2.2021

the visual perspective on a virtual guidance visualisation, especially for handling physical loads, shows the necessity of investigations on:

RQ1: How does the visual perspective on a virtual guidance visualisation influence motor learning in Virtual Reality?

The answer to RQ1 will enable designers of Virtual Reality motor learning training systems to choose a suitable visual perspective for their project based on an empirical basis.

1.1. Outline

This master's thesis aims to increase the empirical evidence of how the visual perspective influences motor learning in Virtual Reality by proposing an experiment. This document will present the development of an experiment and a system with which the experiment can produce this data.
First, the theoretical foundations are laid out, and a state-of-the-art analysis on the basis of related work is provided in chapter 2. With the fundamentals at hand, chapter 3 describes in detailed the design of the proposed experiment. Chapter 3 will adress the independent variables (3.1) and dependent variables (3.3) of the experiment as well the task design (3.2).
The implementation of the proposed experiment is pictured in chapter 4. The experiment requires a digital representation of the learner (4.1) and the guidance visualisation (4.4). Furthermore, physical and digital artefacts the learner will interact with need to be constructed and built physically and digitally (4.2). The digital representations of learner and guidance visualisation as well as the artefacts are then arranged to form the visual perspectives (4.5). Subsequently, the acquisition of the quantitative (4.6) and qualitative (4.7) data is described. The elements of the experiment design and E(x|g)o are composed, and an experiment procedure is defined (4.8). Finally, the limitations of the proposed experiment are pointed out (4.9). The experiment was evaluated by a pilot study, the results and possible improvements are shown in chapter 5. In the end, this master's thesis closes with a conclusion and outlook in chapter 6. The appendix provides a Glossarium, additional figures and the experiment's documents.

2. Motor Learning in Virtual Reality

This chapter provides the theoretical background of Virtual Reality, motor learning, visual perspectives, handling physical load and injury risk metrics in a condensed form. For a more detailed description of these topics, please refer to the preceding seminar thesis [26] which is also digitally attached to this master's thesis. These topics are the essential aspects that serve as the foundation for this master's thesis. Subsequently, an analysis of related work is given. Finally, the research contribution statement is provided.

2.1. Virtual Reality

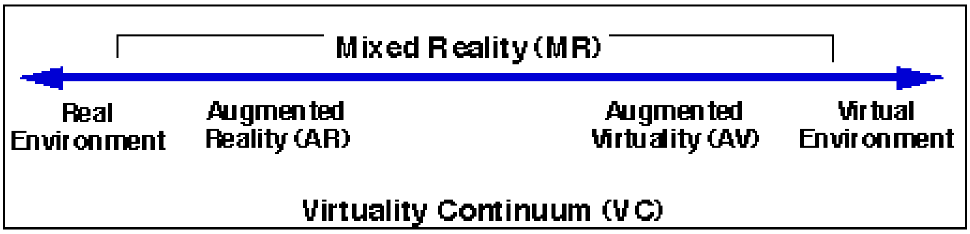


Figure 2.1.: Mixed Reality continuum by Milgram and Kishinho [27]

Milgram and Kishinho [27] describe Mixed Reality (MR) for visual displays on a continuum, compare figure 2.1. Virtual Reality (VR) is purely digital, and thereby the environment is blocked entirely. In Augmented Reality, the environment is visible and augmented with digital elements. During motor learning, the visual perception of one's own body is desirable because it is the most exact representation of one's own body. Thereby, the approach of augmenting the real-world body with a virtual guidance visualisation (GV) is promising. However, today's AR-technology provides a small field of view. A solution to this could be the video see-through technology, but it is limited by latency and distortion [28].

The body's perception can also be achieved by tracking the learner's body and render it over the learner's physical body. Thus, the visual perception of the learner's body can be established in VR. Consequentially, this work will focus on motor learning in Virtual Reality.

2.2. Motor Learning

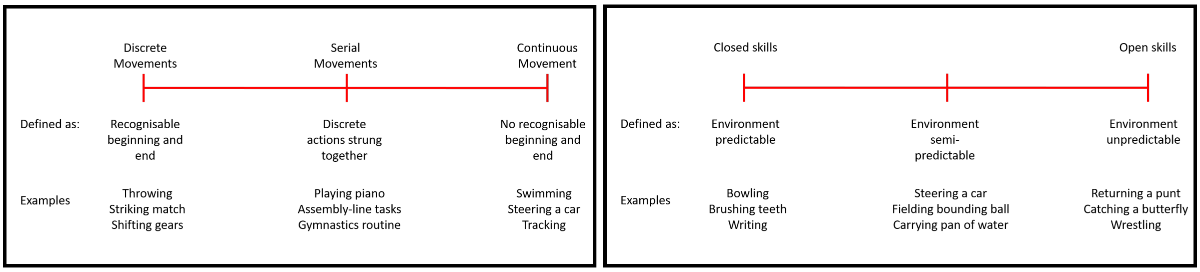


Figure 2.2.: Movement classification by *particualar movement* (left) and *perceptual attributes* by Schmidt et al. [1]

Motor learning is achieved through instruction, attempt, imitation or a combination of them [1]. The process of motor learning can be divided into three parts: cognitive stage, associative stage and autonomous stage (ibid.). In the cognitive stage, training methods are most efficient, and the performance gain is the highest among the stages (ibid.). Tasks that belong to this stage are thereby best suited for an experiment.

Movements can be classified by two means: by *particular movements* and based on *perceptual attributes* (ibid.). Based on the *particular movements*, the classification is described by a continuum, compare figure 2.2 left. On the extremes of the continuum are *discrete movements* and *continuous movements*. Between these extremes, *serial movements* are located. *Discrete movements* are too short for an evaluation. *Continuous movements* do not have a recognisable beginning and end, and thereby they are not suitable for the experiment in question either. *Serial movements* are chained *discrete movements* with a recognisable beginning and end. This allows a task decomposition and an evaluation of particular subtasks. Furthermore, *serial movements* are generally more complex than *discrete* or *continuous movements* and therefore more suited for movement training systems. *Serial movements* are widely used for research in motor learning, for example [17, 18, 23]. Therefore, the experiment task of this master's thesis is based on *discrete movements*.

The classification based on the *perceptual attributes* is also represented by a continuum and includes the environment in which the movement is performed, compare figure 2.2 right. At the extremes of the continuum, *open skills* and *closed skills* are located. For *closed skills*, the environment is predictable, while for *open skills*, the environment is not predictable. The experiment aims to analyse the learner's performance of following a movement and not how they can adapt to environmental changes. Thereby, this experiments task for this master's thesis must be located on the left-hand side of the continuum: *closed skills*.

2.2.1. Measurements for Motor Learning

The movements of a teacher and the movement of a learner differ. To assess the difference between the two movements, two main classes of measures can be applied [1]: *measures of error for a single object* and *measures of time and speed*. *Measure of error for a single object* represent the degree to which the target movement is amiss. Schmidt et al. [1] provide five *error measures* to calculate this error. Among them, *constant error* is the most common measure in related work to determine the difference between the movement of the learner and the movement of the teacher, for example [2, 3, 10, 20, 21, 23, 29]. *Constant Error* is defined as the average error between the learner's movement and the teacher's movement and is described as

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

with x_i: actual value, T: target value, n: number of values [1].

The basic idea of *measures of time and speed* 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. In related work, this measure is mostly assessed by the *task completion time* (TCT), for example [20, 23, 29].

2.3. Visual Perspectives

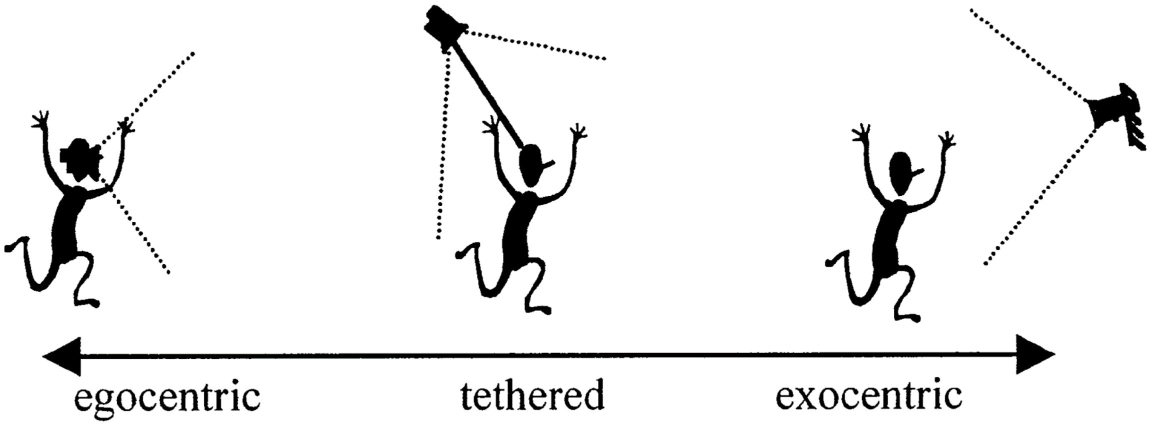


Figure 2.3.: Centricity continuum by Wang and Milgram [30]

Wang and Milgram [30] describe visual perspectives (VP) by the *centricity continuum*, compare figure 2.3. On the left extreme on the continuum, the ego-centric VP is located, in literature also called first-person perspective (1PP). On the right extreme is the exo-centric VP, in literature also called third-person perspective (3PP). The middle part represents tethered VP. By moving from the left to the right, the so-called *tethering distance* increases. The *tethering distance* describes the distance of the anchor point of the eyes to the object in question. In this master's thesis, the object in question is the human-shaped guidance visualisation (avatar). VPs can be clustered into three classes: ego-centric VPs (g-class), exo-centric VPs (x-class) and VPs that contain both ego-centric and exo-centric VPs (gx-class). Without the usage of additional perspective influencing artefacts like mirrors, cameras or screens, there are five possible VPs:

- **Ego-centric**: the teacher's avatar is located inside the learner's avatar. The learner sees the GV inside one's own body, compare figure 2.4 top left and figure 4.9.

- **Purely exo-centric**: the teacher's avatar is located outside the learner's avatar. The learner sees the GV, e.g. in front of him/her, compare figure 2.4 top middle.

- **Augmented exo-centric**: the teacher's avatar is located outside the learner's avatar. Additionally, a virtual copy of the learner's avatar is located inside the teacher's avatar, compare figure 2.4 bottom middle and figure 4.10.

- **Purely ego- & exo-centric**: the combination of purely ego-centric VP and purely exo-centric VP. The learner sees the GV as well as inside and outside of one's own body, compare figure 2.4 top right.

- **Ego- & augmented exo-centric**: the combination of the ego-centric VP and the augmented exo-centric VP. The learner sees the GV inside one's own body, as well as outside. Additionally, a virtual copy of the learner is located inside the exo-centric GV, compare figure 2.4 bottom right and figure 4.11.

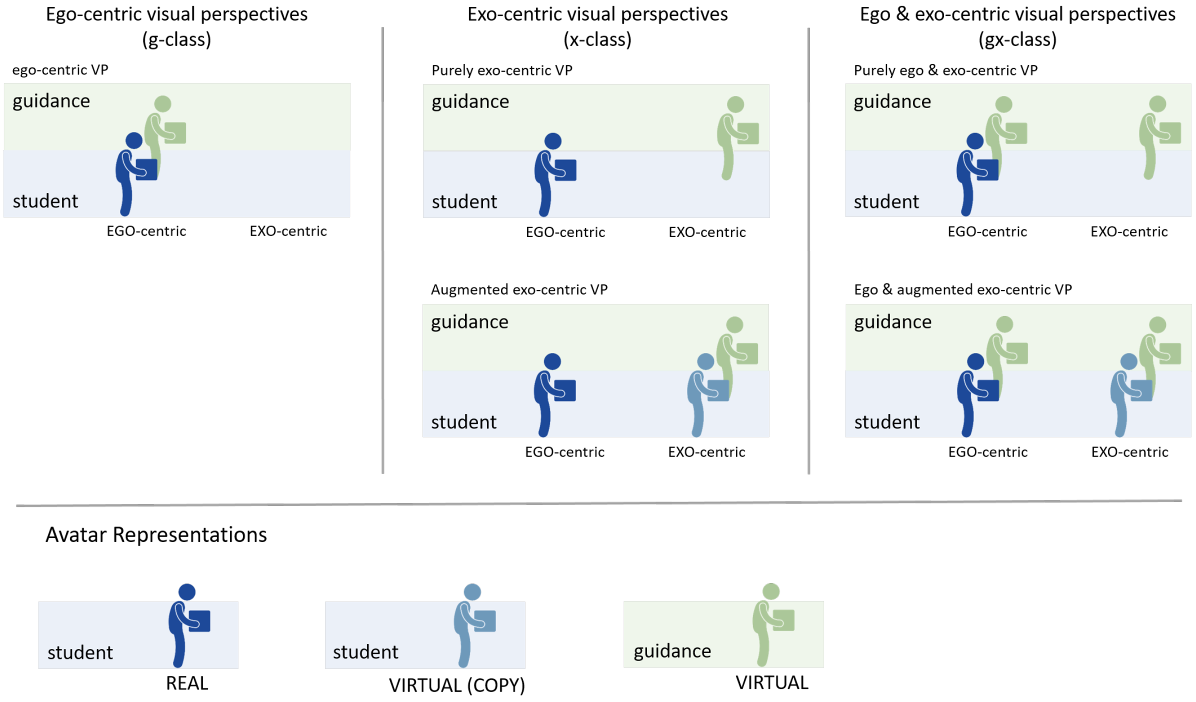


Figure 2.4.: Visual perspectives on virtual guidance visualisations, clustered by their corresponding class. Icon: created by Ghan Khoon Lay from Noun Project, https://thenounproject.com/, accessed: 19.06.2020

2.4. Handling Physical Load

Handling physical load is part of the more general topic Manual Material Handling (MMH). MMH is composed of five elemental tasks: lift, lower, push, pull and hold [31]. Additionally, there are non-elemental tasks like turning and sliding (ibid.). The experiment proposed in this master's thesis will use an experiment task that includes the handling of physical load. Evidently, the task should consist of these elemental tasks. A task that consists of elemental tasks can be generalised to other tasks to a certain extend. To gain a more robust data basis, multiple elemental tasks can be chained together and repeated to form a so-called Unit-Combined-MMH task (ibid.). In chapter 3.2 is described how the elemental tasks become subtasks of the experiment task.

2.5. Injury Risk Metrics

Muckell et al. [32] identified four main features which are common in the bio-mechanical evaluation of different lifting and carrying techniques. Based on those four features, they defined four injury risk metrics to define low risk and high risk movements. The four risk metrics (RM) are described in the following. *Support base* describes the distance between the feet. With a proper support base, an individual is more stable while performing a movement like lifting or lowering. *Squat* describes the distance between pelvis and floor. "A proper *squat* reduces injury risk since the lifting force is applied using legs and not the back" (ibid.). *Upright stance* is defined by the angle between the upright vector and the bend of the back of an individual. *Spine twist* is the angle between the lines between the left and right shoulder and the left and right hip.

2.6. Related Work: Motor Learning in Virtual Reality

g-class	x-class	g-class and x-class	gx-class
AR-Arm [14]	MotionMA [9]	OneBody [20]	Tai Chi Trainer [21]
Just Follow Me [15]	YouMove [2]	LightGuide [23]	SleeveAR [12]
Ghostman [33]	VR Dance Trainer [3]	MR Dance Trainer [6]	
Stylo Handifact [16]	Physio@Home [10]	Throw Simulator [7]	
GhostHands [34]	OutsideMe [4]	Training Phys. Skill [8]	
	E-Learning MA [17]	VP Matters [29]	
	My Tai Chi [18]		
	Perform. Training [5]		
	RT Gestrue Recognition [19]		
	KinoHaptics [11]		
	TIKL [22]		

Table 2.1.: Overview of related work clustered by visual perspectives.

Training movements in Virtual Reality were investigated previously in several works. An overview differentiated by the VP the reasearchers used to train movements is provided in table 2.1. AR-Arm [14], Just Follow Me [15], Ghostman [33], Stylo and Handifact [16] and GhostHands [34] used the ego-centric VP. MotionMA [9], YouMove [2], VR Dance Trainer [3], Physio@Home [10], OutsideMe [4], E-learning Martial Arts [17], My Tai Chi coaches [18], Performance Training [5], Real Time Gesture Recognition [19], KinoHaptics [11] and TIKL [22] used a VP from the g-class. There are also works that used to train movements in g-class and x-class, like OneBody [20], LightGuide [23], Mixed Reality Dance Trainer [6], Free Throw Simulator [7] and Training Physical Skills [8]. It is little research done for movement training in the gx-class, for example Tai Chi Trainer [21] and SleeveAR [12]. The experiment proposed in this master's thesis will compare the ego-centric VP, the augmented exo-centric VP and the ego & augmented exo-centric VP. In the ego & augmented exo-centric VP a virtual copy of the learner is located inside exo-centric GV, which was not part of previous works.

The task which the refferred works use araise from the fields of dancing [2–6], sports [7, 8], rehabilitation [9–13], arts [14–21] and others [22, 23]. None of them include the ergonomic handling of a physical load, but sometimes include physical artefacts like a ball (e.g. Free Throw Simulator [7]) or chop sticks (Ghostman [33]). Also none include locomotion movements. The body parts that are included in the above-mentioned tasks vary, too. For example, [2, 3, 20, 21] full-body movements are taken into consideration, while [12, 14, 34] uses arm movements. The experiment proposed in this master's thesis will utilise a task for full-body movements which, includes the handling of physical load and locomotion movements.

The guidance visualisations which are used to train movements are stick figures [2, 3, 5, 20], wireframes[6, 21], human-shaped avatars[3, 8, 18, 21] and indicators [10, 12, 14, 33]. The experiment proposed in this master's thesis will use human-shaped avatars.

To determine to what extend the movements of the learner matches the GV, different measures are applied. Most common are performance measures based on the accuracy of the performed movements ([2, 3, 10, 20, 21, 23]) and the time related measurements like the *task completion time* ([20, 23]). The experiment proposed in this master's thesis will utilise accuracy measurements, timely measurements, a measurement to assess the learners visual focus and assess quantitative data.

How the perspective influences the learner's performance is sparsely investigated. Recently, in December 2020, Yu et al. [29] conducted three independent studies to close this gap. In the first study, Yu et al. compared the ego-centric VP and a 2D-mirror for single arm movements. In the second study, they compared the ego-centric and exo-centric VP for Yoga. In the third study, they compared the ego-centric VP with a 3D-mirror for arm movements. Yu et al. conclude their findings in a design guideline for systems training motor learning in Virtual Reality: use the ego-centric VP if the type of motion allows, consider alternatives for other types of motions (ibid.). In all three studies, the ego-centric VP outperformed the other perspectives if the movement was completely visible from the ego-centric VP. [20, 23] compared their ego-centric VP with their exo-centric VP. For the task they used, the ego-centric VP outperformed the exo-centric VP. [2, 3] compared the movement learning in VR with traditional video-based movement learning. In both cases, VR movement learning outperformed video movement learning.

2.6.1. Research Contribution Statement

Overall, to my knowledge, motor learning in the gx-class VPs is rarely investigated, especially for full-body movements. Furthermore, motor learning that includes the handling of a physical load in VR in different VPs was not part of investigations. Additionally, previous works used stationary tasks in the ego-centric VP. Moreover, how the visual perspectives influence the learner's visual focus is unexplored.
The proposed experiment in this master's thesis will provide an empirical contribution by increasing the empirical evidence of how the VP on GVs influences motor learning by guiding full-body movements in three VPs: ego-centric, exo-centric and the pure combination of them. Furthermore, empirical evidence can be generated with the proposed experiment for motor learning, including the handling of a physical load and the learner's visual focus. An artefact contribution is provided by presenting a method for guiding locomotion movements in the ego-centric VP.

The generated data of the proposed experiment will help designers of VR motor learning systems to choose a reasonable perspective for their project.

3. Experiment Design

This master's thesis proposes an experiment that answers the research question RQ1: How does the visual perspective on a virtual guidance visualisation influence motor learning in Virtual Reality?

To answer the main research question RQ1, several aspects have to be taken into account: accuracy of movements, transfer of information about how to move, the visual focus of the learner, and last but not least, the personal preference of the learner. Therefore, to answer the main research question RQ1, it is necessary to answer the following sub-research questions:

RQ1.1 How does the visual perspective on a virtual guidance visualisation influence movements' accuracy?

  RQ1.1.1 How does the visual perspective on a virtual guidance visualisation influence movements' accuracy of one's body?

  RQ1.1.2 How does the visual perspective on a virtual guidance visualisation influence the accuracy of handling physical load?

  RQ1.1.3 How does the visual perspective on a virtual guidance visualisation influence the subtasks' accuracy?

RQ1.2 Does the visual perspective on a virtual guidance visualisation influence the transfer of ergonomic principles?

RQ1.3 How does the visual perspective on a virtual guidance visualisation influence the learner's visual focus?

RQ1.4 What is the subjective personal preference of the learner for the visual perspectives?

This chapter describes the design of an experiment which can generate the data to answer the researchquestions RQ1.1-4. First, the independent variables, namely the VPs, are determined in section 3.1. Afterwards, the task for the experiment is developed in section 3.2. Finally, section 3.3 describes the independent variables of the experiment.

3.1. Independent Variables

The last chapter pointed out five VPs, compare figure 2.4. All VPs are worth investigating, and a comparative experiment with all five visual perspectives is desirable. However, to reduce complexity and the number of participants[1], this work will focus on three visual perspectives.
Figure 2.4 shows three main classes of VPs: ego-centric, exo-centric and perspectives which contain both. To answer the research question, it is indispensable to examine at least one of each class. The ego-centric VP is the only VP in the g-class and thus chosen by default. The exo-centric VP can be realised as purely exo-centric or augmented exo-centric. The combination of ego-centric and exo-centric can be realised as ego- & exo-centric or ego- & augmented exo-centric. However, before the exo-centric VP and the combination can be chosen, a closer look at the mechanics that make motor learning in VR possible is necessary.

3.1.1. Excursion: Mechanics for Motor Learning in Virtual Reality

For teaching movements in Virtual Reality in the exo-centric VP, the following issue arises: the GV can move out of the learner's field of view by the movement itself. Szenario: the learner and the GV stand side-by-side. The learner sees the GV to the left. The GV now indicates a movement to turn by 90 degrees to the right. As soon as the learner follows this movement, the GV will move out of the field of view of the learner. After the movement ends, the GV is located behind the learner. The learner cannot see a GV standing behind himself/herself.
This issue is solved in existing work with either the restriction of movements [7, 17] or multiple representations of the GV around the learner [18, 21]. The restriction of movements has a strong influence on the task design and is therefore not desirable for the experiment proposed in this master's thesis. Consequentially, for exo-centric VPs, *multiple representations* of the GVs on strategic positions around the learner are necessary.

In the ego-centric VP, another issue arises during the teaching of locomotion movements. To understand this issue, two aspects have to be clear beforehand: (1) the nature of an ego-centric GV is to be located inside the learner at any time. (2) A GV indicates movements by moving itself. If the GV is about to indicate a movement away from the learner, the GV is moving out of the student's body. However, a GV that is outside of the learner's body is no longer ego-centric.
A possible solution is given by the centricity continuum by Wang and Milgram 2.3. Following the centricity continuum's nature, the tethering distance can be increased by a small amount, and the VP can still be classified as ego-centric. But now the question arises of how far the tethering distance can be increased with which the perspective still feels ego-centric, but the indication of the movement is considerable. For simplicity, this distance is further called *ego-centric tethering distance* (ETD). To determine a reasonable ETD, an informal formative test[2] was conducted with one participant. The participant was a former Computer Science student with expertise in VR systems but had no prior knowledge about

---

[1] During the COVID-19 pandemic, a study is hardly to conduct with a large amount of participant.

[2] A formal study with more participants was not possible because of the COVID-19 pandemic. This holds for all upcoming formative tests.

motor learning. During the formative test, the participant was asked to follow movements in the ego-centric VP. The first movement was conducted with an ETD of 5cm. For the following movements, the ETD was increased by 5cm each. After each movement, the participant was asked about the ability to follow the movements. The subjective assessment of the participant and my observations yielded best for an ETD between 15cm and 30cm. These two values are further called:

$$ETD_{min} = 15cm$$

$$ETD_{max} = 30cm$$

Based on ETD_{min} and ETD_{max} the *speed mechanic* is developed. The *speed mechanic* controls the speed of the playback of the GV. At ETD_{min} and below, the animation plays at normal speed. At ETD_{max} the GV stops. Between ETD_{min} and ETD_{max} the animation speed of the GV is linearly interpolated, compare figure 3.1. The *speed mechanic* was evaluated by an informal formative test with one participant. The participant was a PhD student of Computer Science and had little experience with VR systems, and none in motor learning. The participant's task was to follow the GV in the ego-centric VP. Observations showed that the participant could follow the movement at ease. The opinion of the participant about the *speed mechanic* was very positive ("It did not run away. I had no problem to follow the woman (ed: GV).").
With this excursion, a reasonable decision for the exo-centric VP and the combination can be made.

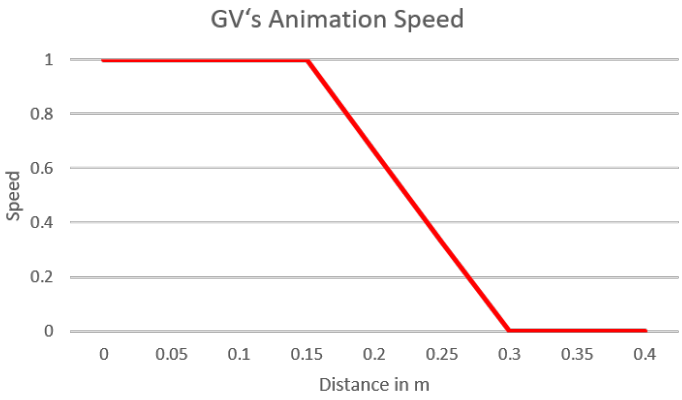
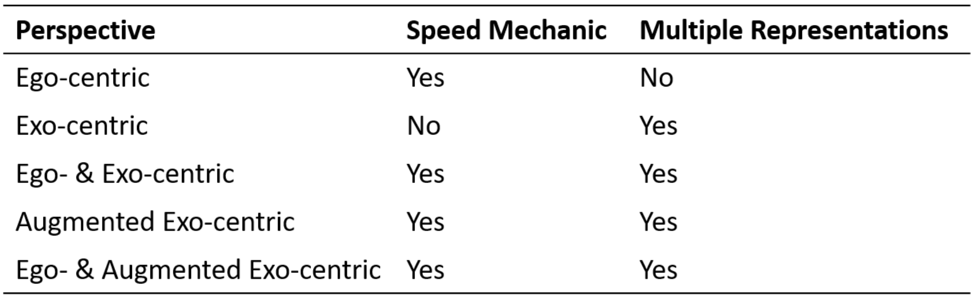


Figure 3.1.: *Speed mechanic*: animation speed of the GV in relation to the distance between learner and GV.

In the ego-centric VP, the learner sees the GV inside the own body. Here, the learner can see the relation of the own body to the GV directly. In the pure exo-centric VP, this relation cannot be seen. Thereby, the position of the learner in relation to the GV must be guessed. That, in turn, makes the application of the *speed mechanic* — which is necessary for ego-centric guidance — nearly impossible. A mechanic that is used in all conditions but one could lead to biased data, compare table 3.1. The mechanic of multiple representations does not influence the experiment's validity because the mechanic would solve an issue that does not exist in the ego-centric perspectives. Furthermore, any VP with more than one



Perspective	Speed Mechanic	Multiple Representations
Ego-centric	Yes	No
Exo-centric	No	Yes
Ego- & Exo-centric	Yes	Yes
Augmented Exo-centric	Yes	Yes
Ego- & Augmented Exo-centric	Yes	Yes

Table 3.1.: Application of *speed mechanic* and *multiple representations* per VP.

representation is an exo-centric VP.

In the augmented exo-centric VP, a virtual copy of the learner is located inside the exo-centric GV. The copy lets the learner see the relation of the own body to the GV. Furthermore, augmenting the exo-centric GV with the learner is widely used and evaluated in related work [2, 21]. Consequently, the augmented exo-centric VP will serve as the exo-centric VP.

With the ego-centric VP and exo-centric VP set, the combination can be determined. In the ego-centric VP, the learner has a direct comparison of the own posture to the GV's posture in the ego-centric VP. In the augmented exo-centric VP, the learner has a direct comparison of the own posture and the GV's posture in the exo-centric VP. For a direct comparison of the own posture and the GVs posture in the ego-centric VP AND the exo-centric VP, the ego- & augmented exo-centric VP is chosen as the combination. The ego- & augmented exo-centric VP is the true combination of ego-centric and augmented exo-centric.

For simplification, the augmented exo-centric VP will be further called exo-centric VP, and the ego- & augmented exo-centric will be further called ego- & exo-centric VP.

The ego-centric VP, exo-centric VP and the ego- & exo-centric VP are the independent variables of the experiment and form the three experiment conditions EGO, EXO, EGO & EXO.

3.2. Task Design

Hornbæk [35] identified three main types of tasks in HCI studies: representative tasks, simple tasks and tasks that use task-specific hypothesis. RQ1 states that the main investigation field is motor learning. Motor learning is strongly related to real-world movements. Evidently, the experiment task is a representative task.

Real-world tasks that include the handling of physical load can found in a wide range of activities. For example, a storekeepers job is to clear a palette of cardboard boxes. This task includes unloading the palette, scaling the boxes, measuring the dimensions of the boxes and finally storing them in a rack. Another example is the work at a grinding machine. The worker takes a slug from a shelf and works on it until the slug becomes a workpiece. After that, the workpiece is carried to a measurement instrument to be verified. There are plenty of other examples, but these two already clarify that tasks which include the handling of physical load consist out of the elemental tasks for *manual material handling*: *lift*, *lower*, *push*, *pull*, *hold*.

The idea for the experiment task is to chain these elemental tasks together to create a Unit-Combined-MMH task that representatively stands for a wide range of tasks that includes the handling of physical load. To achieve this, several aspects have to be taken into consideration: (a) the artefacts with which the learner will interact, (b) a reasonable task decomposition into subtasks and their chaining that allows the investigation of subtasks. Moreover, the experiment needs a (c) structure. (c) will reveal the necessity of three tasks, which have to be (d) equally complex. This section will subsequently discuss (a-d) and propose the task for the experiment.

3.2.1. (a) Artefacts

A task that includes the handling of physical load obviously needs a physical load. In real-world tasks, the physical load can be everything a human can handle. The physical load for this task should fulfil the following criteria. First, the load should have a significant weight, that it is perceived as a load, but at the same time, any healthy person with no previous illnesses can handle it without getting injured. Secondly, the physical load should give enough freedom for interactions. A simple box fulfils the criteria and has a relation to physical loads of real-world tasks like the handling of parcels. With a physical load, the elemental tasks of *lift* and *lower* can be realised by lifting and lowering the box from and to the floor.

Push and *pull* can be realised by pushing and pulling the box on the floor, but it can feel cumbersomely. Moreover, in real-world task pushing and pulling a box is made possible in a more ergonomic height if feasible, not least for security reasons. To support *push* and *pull*, a table is introduced. This table stands representatively, for example, for the grinding machine or a parcel sorting table.

Finally, the transitions between the elemental tasks have to be supported to increase the real-world reference. This is achieved by providing a waypoint. The waypoint is a plate on the floor and helps to bring sense in movements. This plate representatively stands, for example, for a scale or second machine. Walking to a scale or lower a box to the scale on the floor increases the real-world reference more than just an empty place in the room. For simplification, in the following, the addressed waypoint is called scale.

3.2.2. (b) Subtasks

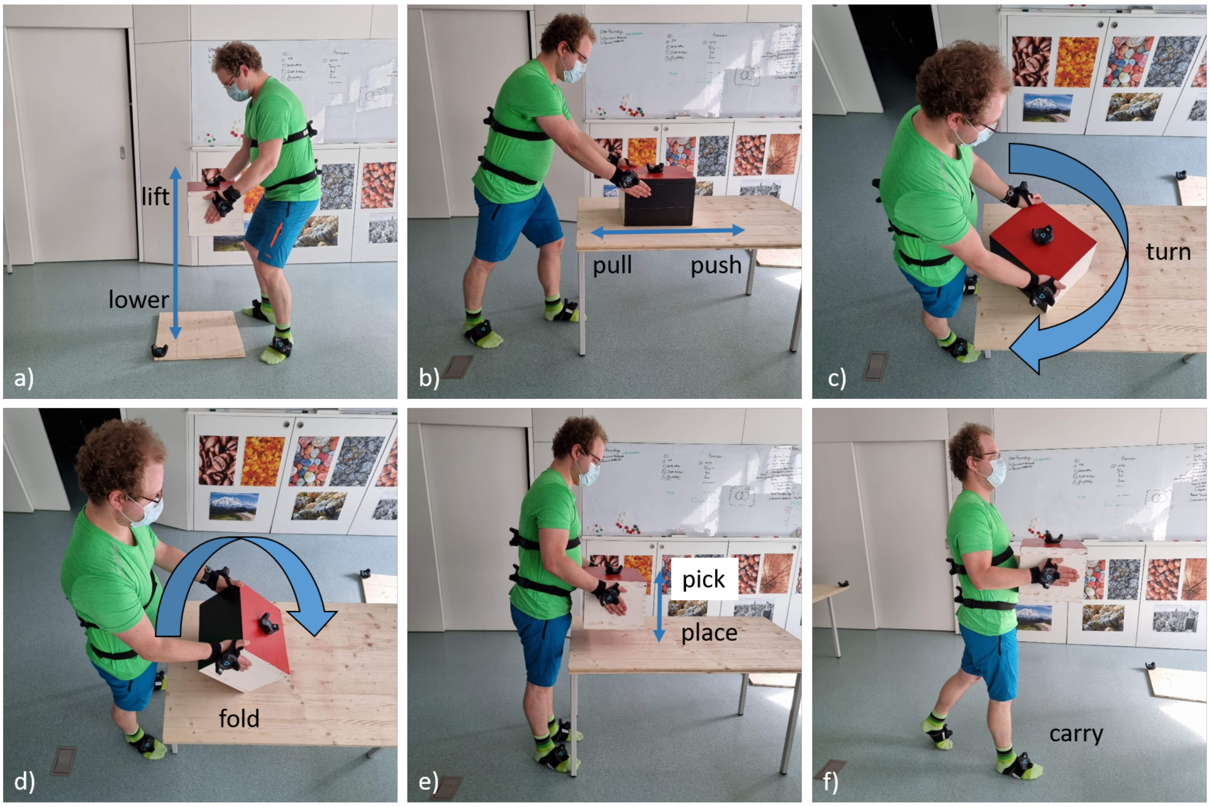


Figure 3.2.: Depiction of subtasks. a) *lift* and *lower*, b) *push* and *pull*, c) *turn*, d) *fold*, e) *pick* and *place*, f) *carry*.

The goal is to create a Unit-Combined-MMH task with the elemental tasks *push*, *pull*, *lift*, *lower* and *hold*. The process of designing the task was complex and took place iteratively. In the following, the process of designing the task is structured by the iterations (task Mk I - task Mk V). For visualisation, the subtasks are depicted in figure 3.2

Task Mk I

The first approach was a task with four occurrences of every elemental task. For lifting the box from the scale and carry the box to the table, obviously, a new task type had to be introduced: *carry*. Because *carry* is not an elemental task and for simplicity, elemental tasks and newly introduced task types are referred to as subtasks. The designing of task Mk I revealed an issue: chaining a given amount of subtasks together so that the task is still conductible is hard to achieve. To overcome the inflexibility in task design, a new subtask is introduced: *walk*. *Walk* means locomotion without the box in hand.

With *walk*, the box can be pushed from one side of the table and then be pushed from the other side of the table, which achieves flexibility in task design. Otherwise, on *push* will always follow *pull*.
Outcome: new subtask *walk* introduced to increase flexibility in task design.

Task Mk II

In task Mk II the subtasks *push*, *pull*, *lift*, *lower*, *carry*, *walk* and *hold* are about to be chained together. Each subtask appeared four times. Task Mk II was informally tested with one participant. The participant had to follow the instructions in the ego-centric VP and exo-centric VP. During the task's conduction, the participant started to look around and correct the own position during the subtask *hold*. An interview afterwards showed that the participant thought the GV had stopped because his position had been too far away from the GV. It became clear that the *speed mechanic* and the subtask *hold* are not compatible. It is indistinguishable for the experiment participant if he/she is too far away from the GV or if it is the subtask *hold*. Because of this indistinguishableness, the subtask *hold* is excluded from the task. However, *hold* is still part of the whole tasks: between the transitions of the tasks (for example, between *lift* ends and *lower* starts) is a slight pause which is equivalent to *hold*. However, this sequence is too short to log reasonably. Furthermore, *hold* is part of the subtask *carry*, where the box is held in front of the body. However, *hold* is not a stand-alone subtask and thus can not be evaluated isolated.
Outcome: subtask *hold* is eliminated because of ambiguity.

Task Mk III

A new task was designed with the subtasks *push*, *pull*, *lift*, *lower*, *carry* and *walk*. During the design, special attention was paid to the magnitude of the movements. For example, every *push* should be equally far. *Lift* and *lower* from and to the scale and *lift* and *lower* from and to the table are very different in magnitude. This resulted in two new subtasks: *pick* and *place*. *Pick* means to pick up the box from the table, *place* means to place the box on the table. For *lift* and *lower* the target remained the scale on the floor.
Outcome: new subtasks *pick* and *place* introduced. This ensures an equal magnitude for every subtask.

Task Mk IV

For task Mk IV the subtasks *push*, *pull*, *lift*, *lower*, *carry*, *walk*, *pick* and *place* are chained together. The task was inspected by a professional physiologist with four years of work experience. The physiologist was asked to describe the subtasks in detail and perform every subtask ergonomically. The professional's description of the subtasks are listed in table 3.2. Through the performance of the physiologist and the description of the subtasks could be derived several insights. The subtasks *push* and *pull* are similar in their conduction. The same applies to the pairs *lift* and *lower* as well as *pick* and *place*. For

the evaluation, this meant that the variations of movements are nearly halved, and thus the possibility of making mistakes is reduced. Example: for *push* and *pull* one foot has to be placed to the back while the other foot remains under the hip. The hands do the same for every *push* and *pull.* If the participant performs the subtask intrinsically correct without the perception of the GV, the experiment will not measure the influence of the perspective.
To increase the number of sources of error, two new subtasks are introduced: *turn* and *fold. Turn* means to turn the box by 90 degrees on the table. *Fold* means to tilt the box from one side to another. The difference in hand movement to *push* and *pull* is obvious. The difference for the feet results from the fact that during *turn* and *fold*, the box' weight remains on the table. The force to apply on the box is significantly lower than during *push* and *pull.* This results in a different feet placement, which is hip wide under the hip.
Outcome: subtasks *turn* and *fold* introduced to increase the possibility of making errors.

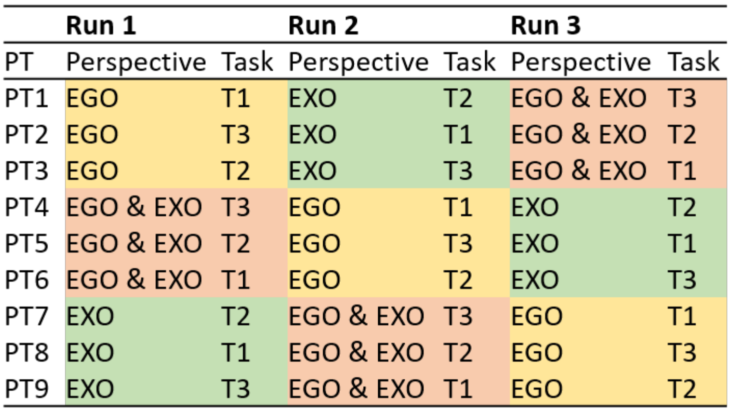
Task Mk V

With the introduction of *turn* and *fold*, all subtasks are introduced. A new task was created with all ten subtasks. To assess all subtasks multiple times, they appear four times each per task. The pair *lift* and *lower* and the only in magnitude different pair *place* and *pick* relate to each other. To be presented equally in the task among the other subtasks, they should also only be present four times. Because *lift* and *lower* are measured with the RM (6.2) (see next section), *lift* and *lower* were decided to appear three times each, and *pick* and *place* one time each. Unfortunately, only one time each *pick* and *place* means that all subtasks that do not happen at the table had to be conducted in sequence. To regain flexibility in the task, it was decided that *pick* and *place* occur two times each. This results in 34 subtasks per task. Table 3.2 provides an overview of the subtasks and their corresponding description, as well as the occurrences per task.
Outcome: task 1 in table 3.3.

Subtask ID	Subtask description	Professional‘s description	#of subtasks per task
push	Push box on table	Lunge, feet hip wide, chest out, shoulders back, straight back, lean forward, bend front knee, extend your arms, pressure on front leg, push box by activating back muscles	4
pull	Pull box on table	Lunge, feet hip wide, chest out, shoulders back, straight back, lean forward, bend front knee, extend your arms, pressure on front leg, pull box by activating back muscles	4
turn	Turn box by 90° on table	Feet hip wide, lean slightly forward with straight back, turn box with arm muscles, weight of the box remains on the table	4
fold	Put the box from one side to another on the table	Feet hip wide, straight back, slightly bended arms, depending on the distance to the box: lean over table, no bent knees, weight of the box remains on the table	4
carry	Translation in space with the box in hand	Chest out, straight back, bend elbows to 90°, box near to body, shoulder in neural-zero	4
walk	Translation in space without the box	“normal walking on their own judgment“, straight back	4
lift	Lift up the box from the floor	Approach box as near as possible, weight shifted slightly to the front, bend knees, open legs while going down, stop at the raised heels, lean forward with straight back, lift box with quadriceps (tights), chest out, elbows aim at ca. 90°	3
lower	Lower box to floor	Head above pelvis, bend knees and open legs, chest out, straight back and head, extend arms	3
place	Put box on table	Parallel hip wide feet, bend knees slightly, lean forward with straight back, lower arms	2
pick	Pick up box from table	Parallel hip wide feet, bend knees slightly, lean forward with straight back, lift with arms, abdominal and back muscles	2
			Total: 34 subtasks per task

Table 3.2.: Subtask ID, description of subtasks and amount occurences of the subtask per task. Professional’s description provided by a professional physiologist.

3.2.3. (c) Experiment Structure



	Run 1		Run 2		Run 3	
PT	Perspective	Task	Perspective	Task	Perspective	Task
PT1	EGO	T1	EXO	T2	EGO & EXO	T3
PT2	EGO	T3	EXO	T1	EGO & EXO	T2
PT3	EGO	T2	EXO	T3	EGO & EXO	T1
PT4	EGO & EXO	T3	EGO	T1	EXO	T2
PT5	EGO & EXO	T2	EGO	T3	EXO	T1
PT6	EGO & EXO	T1	EGO	T2	EXO	T3
PT7	EXO	T2	EGO & EXO	T3	EGO	T1
PT8	EXO	T1	EGO & EXO	T2	EGO	T3
PT9	EXO	T3	EGO & EXO	T1	EGO	T2

Figure 3.3.: Experiment structure: within-subject desing in a Greco-Latin square.

The experiment will compare three conditions: EGO, EXO and EGO & EXO. The main question of this section is how to assign the participants to the independent variables. The key distinction is between within-subject design and between-subject design [35]. In a within-subject design, the participant would experience all conditions. In a between-subject design, the participant would experience only one condition. Within-subject designs typically can detect the differences between the conditions more precisely (ibid.). Furthermore, within-subject designs need fewer participants than between-subject designs (ibid.). For those reasons, the experiment is planned to be conducted in a within-subject design. However, the within-subject design also has a drawback: the participants gain experience about the (i) task and the (ii) conditions during the experiment.

The solution for (i) is to create three tasks with nearly equal complexity. The participant will face a new task every condition. However, the tasks are still similar, and the learning effect persists. A further reduction of the influence of the learning effect on the outcome can be countered out by counterbalancing the task.

(ii) implies that one condition is influenced by another condition, which the participant already experienced. Additionally, there is an asymmetrical carry-over effect between the conditions: EGO & EXO contains condition EGO and EXO[3]. Thereby, EGO & EXO influences EGO and EXO more than EGO and EXO influence EGO & EXO. The solution to (ii) is counterbalancing, to counter the effect out.

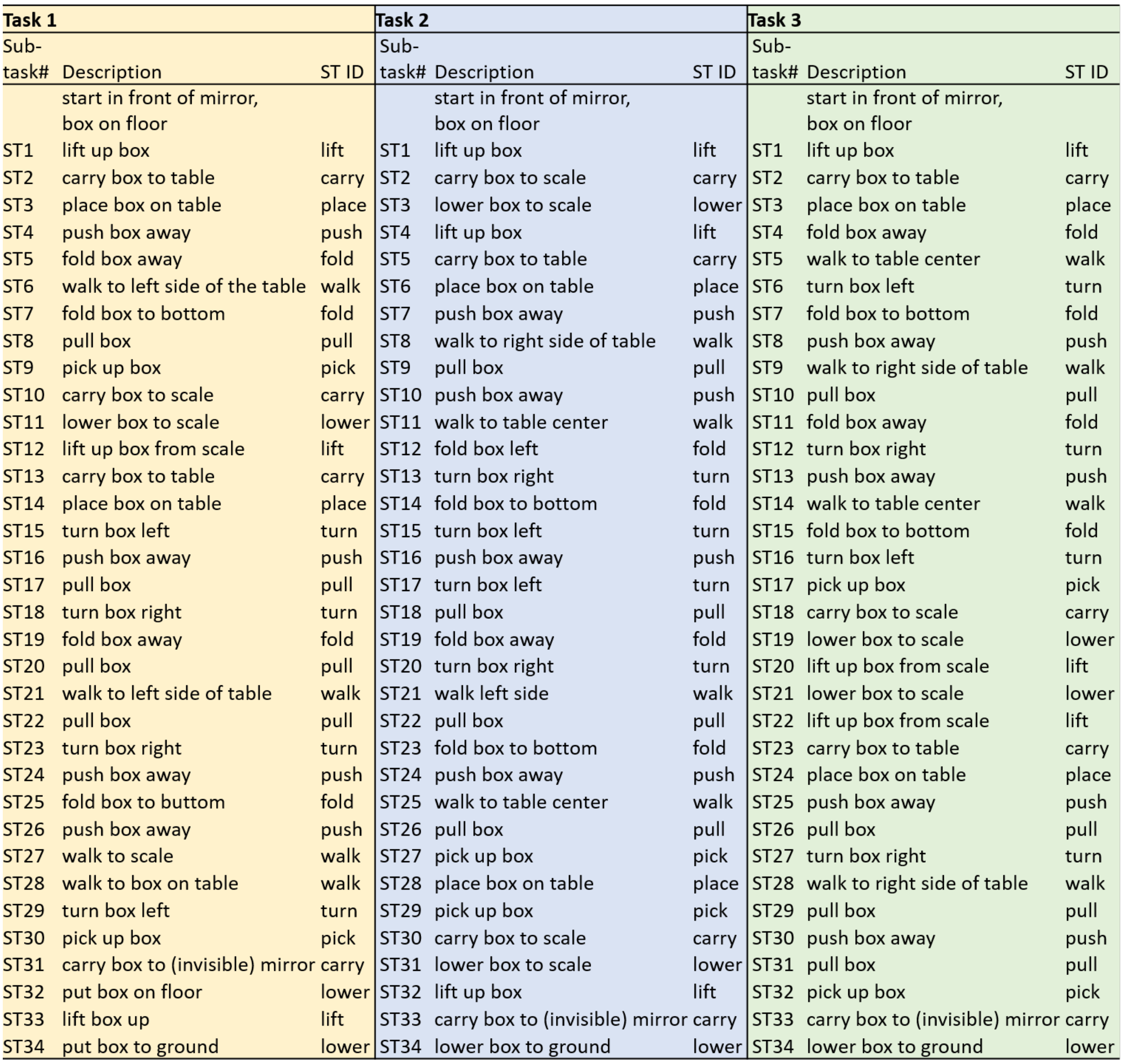
Hornbæk proposes, in this case, to cross the conditions with the task and use a Greco-Latin square [35]. Three conditions and three tasks in a Greco-Latin square result in blocks of nine participants. A block is depicted in figure 3.3. Apart from this, Hornbæk states that experiments conducted within-subject

---

[3] EGO & EXO is the union of EGO and EXO

should be conducted with at least 20 participants (ibid.). Because one block requires nine participants, the experiment should be conducted with at least three blocks ($3x9 = 27$ participants). The participants will have different demographics, which can influence the experiment's outcome, too. To reduce the demographic effect, the first *run* (the first task performed by the participant) of every participant is for acclimatisation and is excluded from evaluation.

3.2.4. (d) Equal Task Complexity



Task 1			**Task 2**			**Task 3**		
Sub-task#	Description	ST ID	Sub-task#	Description	ST ID	Sub-task#	Description	ST ID
	start in front of mirror, box on floor			start in front of mirror, box on floor			start in front of mirror, box on floor	
ST1	lift up box	lift	ST1	lift up box	lift	ST1	lift up box	lift
ST2	carry box to table	carry	ST2	carry box to scale	carry	ST2	carry box to table	carry
ST3	place box on table	place	ST3	lower box to scale	lower	ST3	place box on table	place
ST4	push box away	push	ST4	lift up box	lift	ST4	fold box away	fold
ST5	fold box away	fold	ST5	carry box to table	carry	ST5	walk to table center	walk
ST6	walk to left side of the table	walk	ST6	place box on table	place	ST6	turn box left	turn
ST7	fold box to bottom	fold	ST7	push box away	push	ST7	fold box to bottom	fold
ST8	pull box	pull	ST8	walk to right side of table	walk	ST8	push box away	push
ST9	pick up box	pick	ST9	pull box	pull	ST9	walk to right side of table	walk
ST10	carry box to scale	carry	ST10	push box away	push	ST10	pull box	pull
ST11	lower box to scale	lower	ST11	walk to table center	walk	ST11	fold box away	fold
ST12	lift up box from scale	lift	ST12	fold box left	fold	ST12	turn box right	turn
ST13	carry box to table	carry	ST13	turn box right	turn	ST13	push box away	push
ST14	place box on table	place	ST14	fold box to bottom	fold	ST14	walk to table center	walk
ST15	turn box left	turn	ST15	turn box left	turn	ST15	fold box to bottom	fold
ST16	push box away	push	ST16	push box away	push	ST16	turn box left	turn
ST17	pull box	pull	ST17	turn box left	turn	ST17	pick up box	pick
ST18	turn box right	turn	ST18	pull box	pull	ST18	carry box to scale	carry
ST19	fold box away	fold	ST19	fold box away	fold	ST19	lower box to scale	lower
ST20	pull box	pull	ST20	turn box right	turn	ST20	lift up box from scale	lift
ST21	walk to left side of table	walk	ST21	walk left side	walk	ST21	lower box to scale	lower
ST22	pull box	pull	ST22	pull box	pull	ST22	lift up box from scale	lift
ST23	turn box right	turn	ST23	fold box to bottom	fold	ST23	carry box to table	carry
ST24	push box away	push	ST24	push box away	push	ST24	place box on table	place
ST25	fold box to buttom	fold	ST25	walk to table center	walk	ST25	push box away	push
ST26	push box away	push	ST26	pull box	pull	ST26	pull box	pull
ST27	walk to scale	walk	ST27	pick up box	pick	ST27	turn box right	turn
ST28	walk to box on table	walk	ST28	place box on table	place	ST28	walk to right side of table	walk
ST29	turn box left	turn	ST29	pick up box	pick	ST29	pull box	pull
ST30	pick up box	pick	ST30	carry box to scale	carry	ST30	push box away	push
ST31	carry box to (invisible) mirror	carry	ST31	lower box to scale	lower	ST31	pull box	pull
ST32	put box on floor	lower	ST32	lift up box	lift	ST32	pick up box	pick
ST33	lift box up	lift	ST33	carry box to (invisible) mirror	carry	ST33	carry box to (invisible) mirror	carry
ST34	put box to ground	lower	ST34	lower box to ground	lower	ST34	lower box to ground	lower

Table 3.3.: Task 1 - 3. ST#: subtask number, ST ID: subtask ID. Reading the description of one task from ST1 to ST34 corresponds to the instructions the learner receives from the GV.

An experiment participant will face in every condition another task. For the experiment's validity, it is indispensable that these three tasks have nearly equal complexity. As described in (b), a task consists of ten subtasks that occur with a specific amount. The main idea to ensure a comparable complexity is to use the subtasks for all three tasks in an equal amount but shuffled. This means the 34 subtasks of task one occur in task two and three but in a different order. Table 3.3 lists all three tasks. For every task, the subtask number ST1-ST34 is provided. Every subtask number stands for a subtask, which comes with a description and the subtask ID. Reading the description from top to bottom is the instruction the learner receives from the GV during one condition. The mirror mentioned in the first line is another waypoint, which is necessary for technical reasons and is described in section 4.1.

3.3. Dependent Variables

This master's thesis aims to answer the main research question RQ1: How does the visual perspective on a virtual guidance visualisation influence motor learning in Virtual Reality? To answer RQ1, the proposed experiment has to generate data that can answer the sub-research questions RQ1.1-4. This section will provide the underlying paradigm to every sub-research question and explain which measures are necessary.

RQ1.1 How does the visual perspective on a virtual guidance visualisation influence movements' accuracy?
Paradigma: The more exact the learner's movements matches the GV movements, the better the learner could follow the instructions of the GV.
For RQ1.1.1, the limbs of the learner and the limbs of the GV are compared. For RQ1.1.2, the box' accuracy is compared. For RQ1.1.3, both are compared and additionally, the current subtask is taken into consideration. The accuracy can indicate how the particular movement is suited for the VP.

> **RQ1.1.1** How does the visual perspective on a virtual guidance visualisation influence movements' accuracy of one's body?
> **Measures**: (1) Euclidean distance between the learner's and GV's hands, feet, head and hip in meters. (2) Angle between the learner's and GV's hands, feet and hip in degrees.
>
> **RQ1.1.2** How does the visual perspective on a virtual guidance visualisation influence the accuracy of handling physical load?
> **Measures**: (3) Euclidean distance between the learner's and GV's box. (4) Angle between the learner's and GV's box in degrees.
>
> **RQ1.1.3**How does the visual perspective on a virtual guidance visualisation influence the subtasks' accuracy?
> **Measures**: (1-4), additionally matched to the subtask that is currently performed (5).

(1-4) gives insights to what extent the learner could follow the GV for the whole task. (5) can extract specific subtasks for which the learner could follow the GV to a certain extent. For example, in the ego-centric VP, the overall accuracy for a task is lower than in the other VPs, but the accuracy for the subtasks *lift* and *lower* is higher than in other VPs. For this example, measure (5) can extract specific subtasks that are performed better or worse than in other VPs.

RQ1.2 Does the visual perspective on a virtual guidance visualisation influence the transfer of ergonomic principles?
Paradigma: the better the RM score, the better the ergonomic principles could be transferred.
Measures: (6) risk metrics: (6.1) *upright stance* in degrees, (6.2) *squat* in meters, (6.3) *support base* in meters, (6.4) box-near-body in meters.

> (6.1) *upright stance* is defined by the difference in degrees between the straight upward vector and the back of the learner. For all subtask, *upright stance* should be in a certain window, see section 4.6.2. *Upright stance* indicates if the learner could percept the correct posture of his back.
>
> (6.2) *squat* is defined by the distance in meters between the feet. For the subtasks *lift* and *lower*, the squat distance should be in a specific window. For the other subtasks, *squat* is not applied because the knees do not bend in the other subtasks. *Squat* indicates if the learner could percept that he/she should bend the knees during *lift* and *lower*.
>
> (6.3) *support base* is defined by the distance in meters between the feet. For the subtasks *push*, *pull*, *turn*, *fold*, *lift*, *lower*, *pick* and *place*, *support base* should be in a specific window. Support base indicates if the learner could percept the correct posture of the feet. Muckell et al.[32] additionally use the RM *spine twist* in their work. This RM cannot be applied for this experiment because of the *multiple representations* mechanic. The learner has multiple GV around and is free of choice which one to look at. The turn of the head implies spine twist. Thus, *spine twist* would have low validity and reliability.
>
> (6.4) *box-near-body*. During the task design, a professional physiologist was consulted. During the interview, all movements were described in detail, compare table 3.2. During the subtask *carry*, the box should be as near as possible to the body, while the elbows should have a bend angle of 90 degrees. The physiologist stated this posture as important. This statement is the basis for introducing an additional risk metric related measurement: *box-near-body*. Unfortunately, the bend angle could not be determined during an experiment for technical reasons, see chapter 4. Fortunately, the distance between the box and hip can be determined. *Box-near-body* is defined as the distance in meters between the learner's box and hip. For the subtask *carry*, *box-near-body* should be in a certain window.

RM (6.1-4) are different from accuracy measurements (1-5) because they are independent of the learner's position and the GVs position. For example, in the exo-centric VP, a learner cannot percept the correct position where he/she should stand. The learner thereby stands 15cm away from the position he/she should stand. The overall accuracy is thereby lower. But the learner could percept the positioning of his/her feet correctly. In this case, the RM (6.3) high while the accuracy is low.

RQ1.3 How does the visual perspective on a virtual guidance visualisation influence the learner's visual focus?
Measures: (7) *looking at*
Paradigma: the learner's visual focus is on the object the learner is looking at.
The learner interacts with a box and has multiple GVs around and inside oneself. *Looking at* can give insights on which GV the learner is focusing, the frequency of focus changes and the role of the physical load.

RQ1.4 What is the subjective personal preference of the learner for the visual perspectives?
Measures: (8) qualitative data; Likert scales, semi-structured interview, digging into incidences. After each *run*, an *questionaire* is handed to the participant. After all three *runs*, a semi-structured interview is conducted.
The qualitative data serves not only to investigate the learner's personal preference but also to serve as triangulation method for (1-7).

The last measure is the (9) task completion time (TCT) measured in milliseconds. The *speed mechanic* regulates the speed of the animation of the GV. The further the learner is located to the GV, the slower the GV animation speed until it stops entirely at EDT_{max}. The task completion time can give insights into what extent the learner could perceive the desired position of one's body in relation to the GV. This measure relates to (1) and can be used for triangulation. Additionally, it is to expect that the TCT is decreasing from condition to condition because the participant acclimates. By that, TCT could give insights into the learning effect between the conditions.
Finally, the experiment is recorded by video. If during the evaluation questions about a specific topic arise, the recordings can be consulted.

4. Experiment Implementation

The previous chapter describes an experiment to investigate the influence of the visual perspective on motor learning in Virtual Reality. For the conduction of this experiment, a system is necessary. This system is called E(x|g)o. For an in-detail description of the implementation of E(x|g)o the preceding project report to this master's thesis can be consulted [36] which is also digitally attached to this document. This chapter elucidates the development of E(x|g)o in a condensed form. The starting point is the creation of a self-perception of the learner. Section 4.1 describes how the learner receives a digital body (avatar) in Virtual Reality. In section 4.2 the physical and digital artefacts are added with which the learner will interact and where they are located in the physical and virtual space (4.3). Then, the GVs are added in section 4.4. Subsequently, section 4.5 describes the implementation of the VPs, which serve as conditions of the experiment. With the learners avatar, artefacts, GVs and the experiment conditions implemented, E(x|g)o is able to teach motor learning in VR. To measure the performance of the experiment participants, the measures from section 3.3 are implemented and described in section 4.6 and 4.7. After E(x|g)o is complete, all actions to perform the experiment are known. These actions are assembled with the experiment design, and the experiment's procedure can be described. The experiment procedure is depicted in section 4.8. Finally, section 4.9 evinces the limitations of E(x|g)o.

4.1. Self-Perception

Figure 4.1.: Hardware utilised by E(x|g)o. a) Valve Index, b) base station, c) Vive Tracker 2, d) Vive Tracker 2 attached to Vive Tracker straps.

There are various options of devices to dive into Virtual Reality. Several devices have been evaluated, and the decision was made for the Valve Index[1] (figure 4.1 a), because of its refresh rate, screen solution, field of view and the possibility to wear glasses underneath the head-mounted display (HMD). To determine the position and orientation of the HMD, the so-called Lighthouse is utilised. A Lighthouse consists of at least two base stations[2] (figure 4.1 b). The base stations are placed at opposite corners of a room and span the tracking volume. To improve the tracking and for the avoidance of untracked areas, e.g. under the table, E(x|g)o uses four base stations, one for each corner of the room, to span the

---

[1] https://www.valvesoftware.com/en/index, accessed 29.3.2021

[2] https://www.valvesoftware.com/en/index/base-stations, accessed 29.3.2021

Lighthouse.
With this setup, the learner can move in an empty virtual world. The next step is to replace the empty virtual world with a meaningful environment. To create the environment, the game engine Unity 3D[3] was used. In Unity3D, a basic room was created. Four light yellow walls, a parquet floor and unidirectional lighting. The parquet floor serves a purpose: it has a structure with frequent straight lines, making it easier to align the artefacts the learner will interact with. The room is kept simple not to distract the participant from the experiment.
Because high realistic GVs tend to perform better than abstract GVs [37] and indicator-based GVs for full-body movements tend to overwhelm learners [23], the decision was made to used high realistic human-shaped avatars for the learner and the GV. The next step is to add the learner's avatar to the empty room. To achieve this, the learner's body needs to be tracked. Multiple full-body tracking systems were compared. The decision was made for Vive Tracker 2[4] (figure 4.1 c), because of the cease of coordinate system matching, low latency and less work-intensive calibration process. The learner wears six Vive Trackers in total, compare figure 4.2. Five of them plus the HMD are necessary for the full-body tracking of the learner. The remainder is necessary for RM (6.1), which is later explained in section 4.6. Two trackers are located at Dorsum pedis[5] (compare figure 4.2 a), two trackers are located at Dorsum manus[5] (compare figure 4.2 b). One tracker is located at Vertebra lumablis 5[5] (L5) (compare figure 4.2 c). The trackers are attached to the learner by special Vive Tracker Straps[6] (figure 4.1 d).
The Lighthouse tracks the Vive Trackers and HMD, which send their position to the PC. On the PC, SteamVR[7] receives the information and forwards it to Unity3D. In Unity3D, the SteamVR Plugin[8] provides the information in a workable format. The tracking information is now about to be transformed into a rendering of a human-like avatar at the position of the learner's body. This requires several steps. First, an avatar is imported. To create the avatar, Reallusion Character Creator 3[9] was used. To match the gender of the participant, a male and a female character was created, wearing the same clothes. Based on the demographic questionnaire, the gender can be set, and the participant will see an avatar complying with the participant's gender.
Secondly, the tracker's position and orientation in the tracking volume have to be translated into humanoid movements that meet the learner's movements. This is achieved by Inverse Kinematics (IK).
Short excursion: IK arises from the field of robotics. A robot arm consists of limbs and joints. Each limb has a specific length, and each joint has a specific range of angles to move. The length and angles are called rules. Given an endpoint the robot has to reach with the most outer limb, the angle of each joint can be calculated with the rules. This process can be mapped to a human body, too.

Unity3D provides a third-person plugin called FinalIK[10] that is capable of the calculations in question. On the one hand, FinalIK is powerful and unrivalled in functionality compared to other IK tools and thus influenced the choice to use Unity3D for E(x|g)oȮn the other hand, to match the needs of the experiment, extensive adjustments were necessary. The main task is to transfer the information from

---

[3] https://unity.com/, accessed 29.3.2021

[4] https://www.vive.com/eu/accessory/vive-tracker/, accessed 29.3.2021

[5] Latin description: Dr. med. univ. Kilian Roth

[6] https://www.google.com/search?q=vive+tracker+straps, accessed 10.03.2021

[7] https://store.steampowered.com/app/250820/SteamVR/, accessed 21.03.2021

[8] https://assetstore.unity.com/packages/tools/integration/steamvr-plugin-32647

[9] https://www.reallusion.com/character-creator/, accessed 21.3.2021

[10] https://assetstore.unity.com/packages/tools/animation/final-ik-14290, accessed 21.03.2021

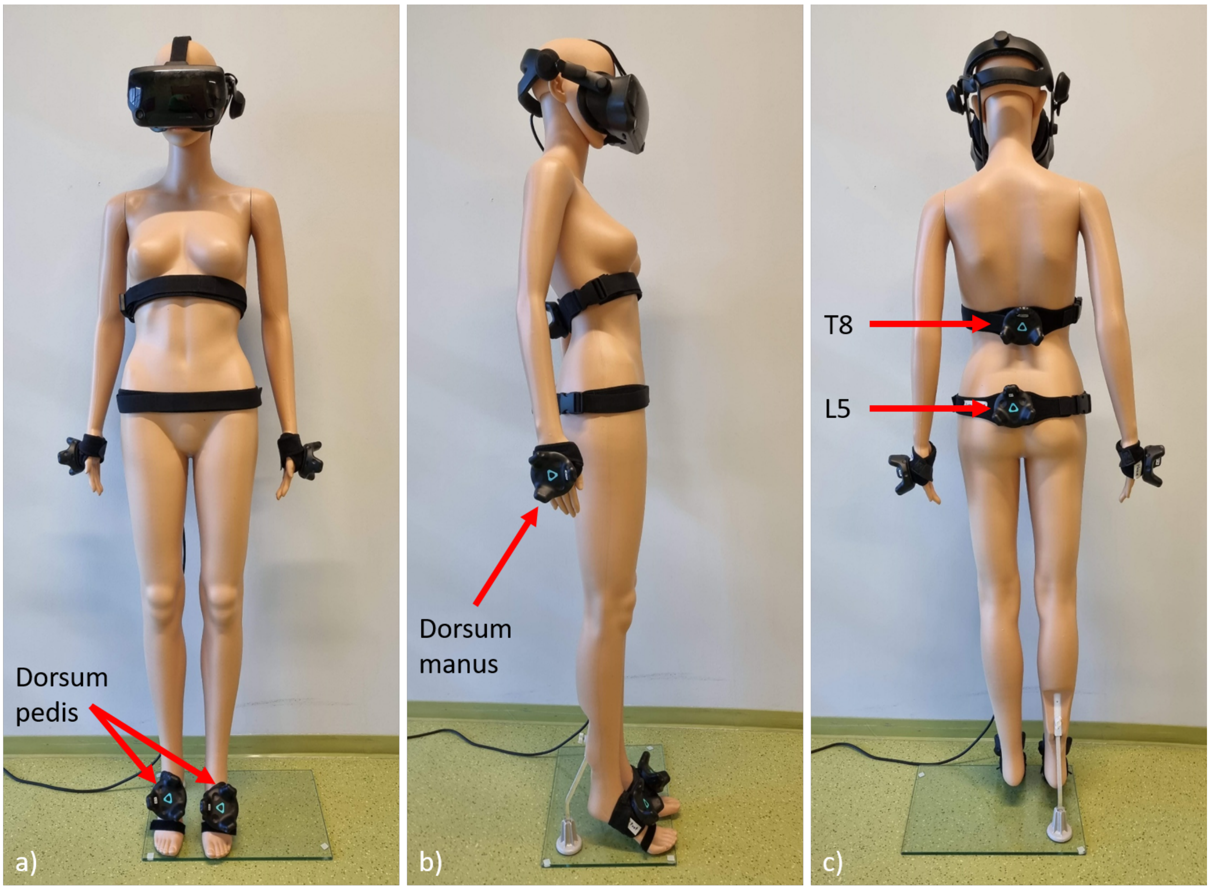


Figure 4.2.: Tracker placement. a) front view - Vive Tracker at Dorsum pedis, b) side view - Vive Tracker at Dorsum manus, c) back view - Vive Tracker at Vertebra lumablis 5 (L5) and Vertebra thoracalis (T8).

SteamVR to FinalIK in a meaningful way so that FinalIK animates the learner's body faithfully.
SteamVR registers the Vive Tracker in the order they are switched on. To increase the reliability of E(x|g)o, a script was created that assigns the tracker by the hardware ID. The trackers are then assigned to a script called VRIKCalibrationController. The VRIKCalibrationController matches the tracker with the avatar and resizes the avatar to the learner's height. FinalIK is constructed to work with controllers in the user's hands. In E(x|g)o, the experiment participants needs the hands to interact with the box, thus the controllers are replaced with Vive Tracker on the back of the hands. Shifting the reference points of the hands yields a faithful representation of the learner's hands. The feet needed similar adjustments. Finally, FinalIK is able to solve the movements. Solving is the process of translating the tracker information into an animated avatar. For clarification, the complete rendering pipeline exemplary for the hip of the learner is attached in appendix B.1.
FinalIK requires calibration before use. For calibration, the person equipped with the trackers needs to perform a T-pose facing a specific direction. To ease the calibration process, a virtual mirror is placed in the room. The participant can be asked to look into the mirror and expand the arms, leading to the participant's correct orientation during the calibration process. Immediate with the calibration, the

mirror disappears.
After the calibration, the system is ready to start with the task. Because the participant is now standing in front of the mirror, the position in front of the mirror is chosen as the starting point and endpoint of every task.
With the steps implemented in this section, the outcome results in a faithful representation of the learner, see figure 4.3.

The approach of translating a real-world person to the digital world with the help of Vive Trackers and IK was used by other researchers, too, for example [29, 38].

Figure 4.3.: Digital representation of real-world person with E(x|g)o. a) real-world person, b) digital avatar representation in VR of the real-world person. Images not taken at the exact same time.

4.2. Physical and Digital Artefacts

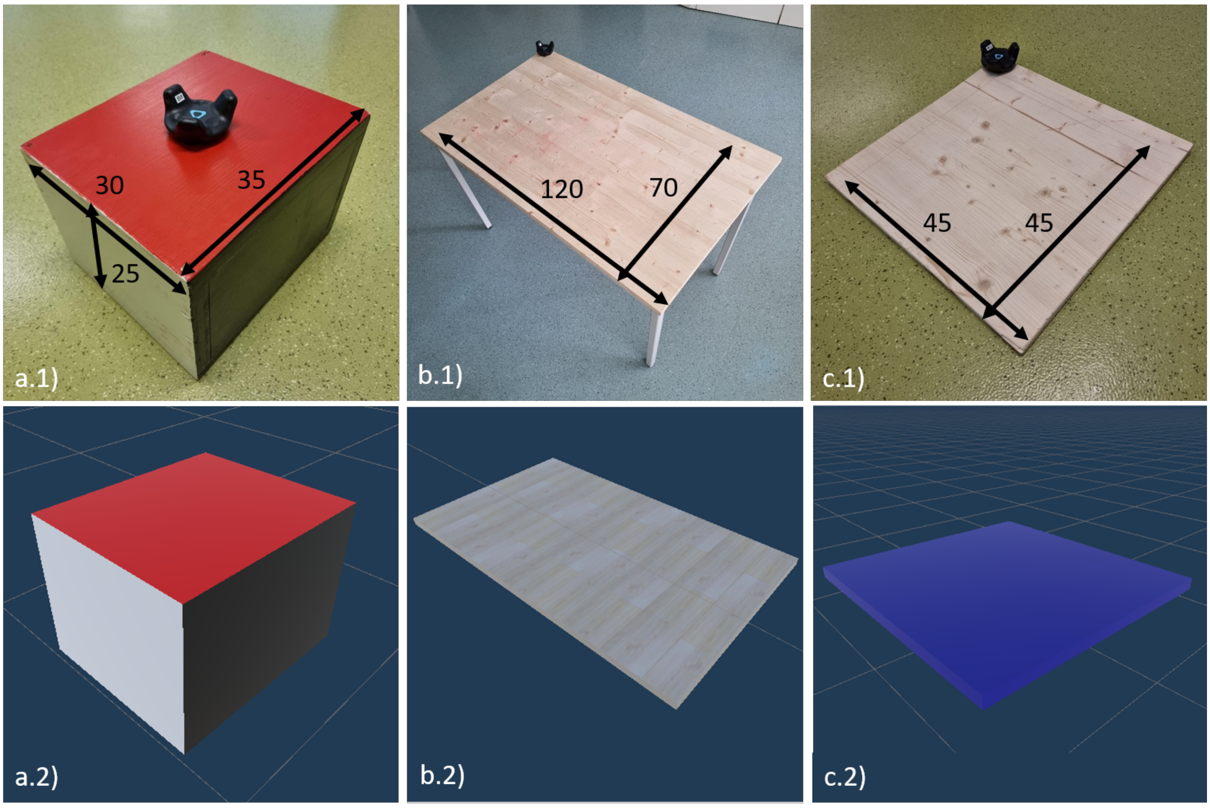


Figure 4.4.: Physical and corresponding digital artefacts of E(x|g)o including measures in cm. a) box, b), table — measures do not include the additional 6cm to attach the tracker, c) scale — measures do not include the additional 5cm to attach the tracker.

With the previously described elements in place, the learner can see the own body in an empty room. The task includes the handling of physical load on a table and a scale. The creation of the table, scale and the box, which will serve as physical load, starts with the construction of them - physically and digitally. Figure 4.4 shows the physical and digital versions of the box (a), table (b) and scale (c).
The first version of E(x|g)o used a cardboard box (27cm x 26cm x 24cm) as physical load. During the development, it became clear that the size of thecardboard box was too small and light to serve as a physical load. To determine a suitable size, several boxes of different dimensions were informally tested. With nine different boxes, a set of subtasks were performed. The major insight from this test was that the length of the box' sides should be unequal to see the direction of the box visually. Furthermore, the box should be perceived as physical load by having a reasonable size and a certain weight. Simultaneously, the box should not be too heavy and thereby limit the experiment participants to strong humans or being a threat to the participants' bodies. The sizes were set for 35cm x 30cm x 25cm. The measures of the box differ by 5cm in every dimension. This makes it clear to see the orientation of the box. The final box was constructed from three-layered wood with a strength of 27mm. The resulting weight was

5.8kg. To evaluate the weight of the box, one male participant[11] was asked to perform all subtasks. The participant is a computer science PhD student and experienced with several sports activities. Observations revealed no incidences that contradict to use the box as physical load: the box could be safely held in hands, and it was visible that the person changed the own posture during the handling of the box to perform the movements more ergonomically. The posture change can be interpreted as a sign that the box is perceived as "load". The person rated the weight as "ok". He had no problems moving the box.

The box was painted in three high contrast colours: black, white and red. Each opposite side was painted in the same colour. The painting facilitates the visual perception of the orientation of the box. The digital pendant of the box is a cuboid in the same colour and size. To translate the physical box' position and orientation to the virtual one, a Vive Tracker is attached to the box and fixated with a screw. The plus side of using a screw is the prevention of any relocation of the tracker. The downside is that tremors caused by placing the box on, for example, the table, are transferred directly into the tracker. This causes the tracker to lose tracking. To interrupt the transfer of tremors, shock-absorbing insulation is placed between the tracker and the box.

During the development of E(x|g)o an office table (120cm x 60cm x 72cm) was used. The digital pendant to the physical table is a plate in the same size and colour as the tabletop. The position and orientation of the table are assessed by a Vive Tracker. Unfortunately, the tracker was placed on the top of the table inside the working area, where the box will be placed and shifted during the tasks. To shift the tracker out of the working area, a new tabletop was constructed. Because the used office table was too narrow, the width was increased by 10 cm. The new tabletop is out of three-layered wood, with an additional increased size of 6cm in length and width (126cm x 76cm instead of 120cm x 60cm) to provide additional space for the tracker. The tracker is attached on the most outer edge, thus out of the working area. To prevent tremors from passing from the table into the tracker, shock-absorbing insulation is applied between the tracker and tabletop.

The last artefact to create is the scale. The scale is a waypoint in the room where the participants perform *lift* and *lower* to the ground. The scale is a rectangular plate of 45cm x 45cm so that the box can be placed on the scale easily. To shift the tracker out of the area where the box will be placed, the plate is extended by 5cm. The tracker is attached to the most outer edge of the extension with a screw and shock-absorbing insulation. The digital pendant is a plate with the exact dimensions of the physical plate, excluding the extended area. The physical artefacts were constructed and built by myself.

---

[11] Evaluation with at least one male and one female person is desirable. This was not possible because of the COVID-19 pandemic.

4.3. Experiment Setting

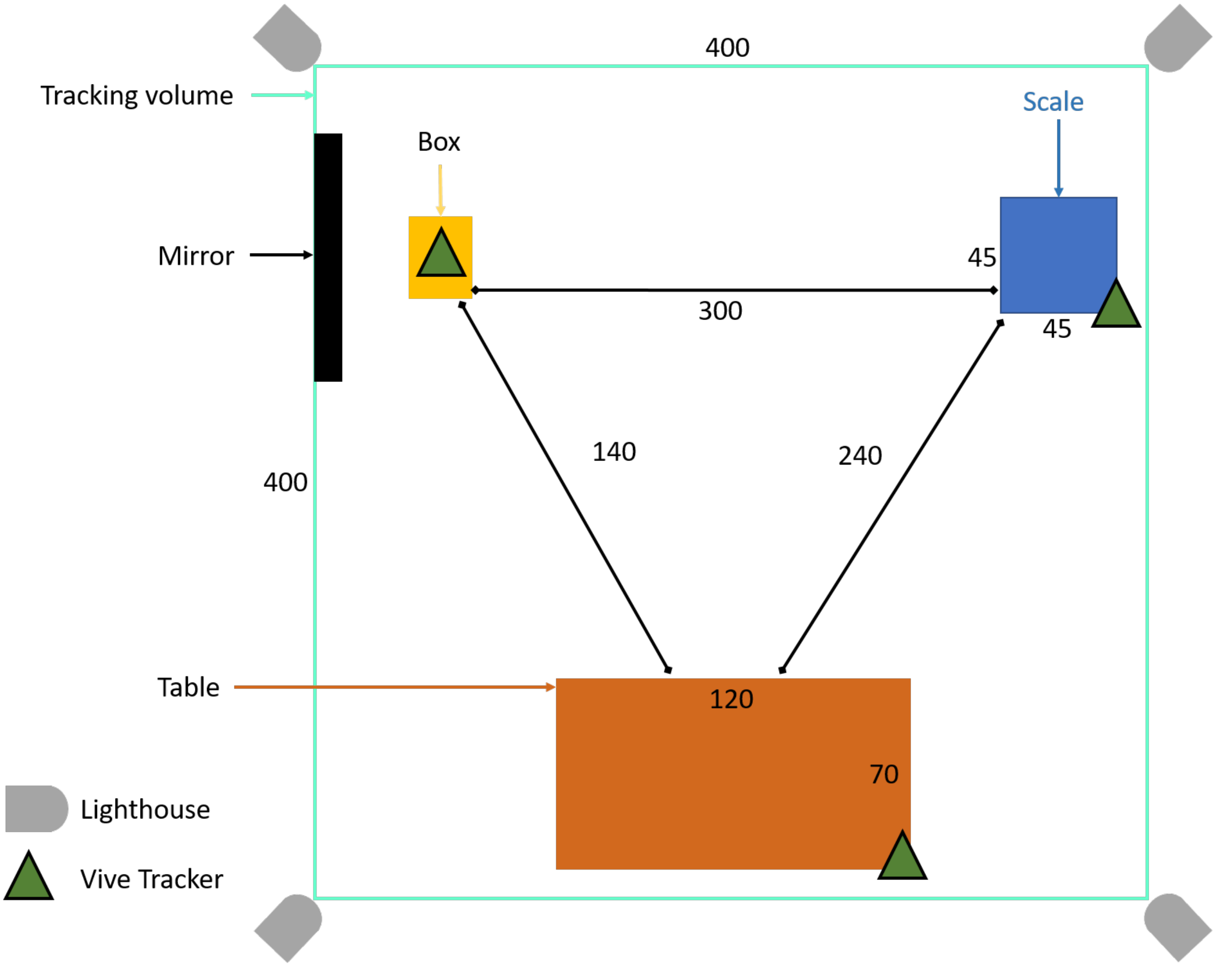


Figure 4.5.: Experiment setting with measures in cm.

.

Meanwhile, E(x|g)o consists of a room, an avatar representing the learner, table, box and scale. In the following, an overview of the alignment of these elements is given. Figure 4.5 shows the sylised virtual room in which these elements are aligned. Figure 4.6 show the corresponsing real-world room. All positions that are described in the task description (compare table 3.3) are depicted. The outer line represents the Lighthouse or tracking volume, which is approximately 400cmx400cm.
On the left wall, the mirror is located. In front of the mirror, the starting and end position of the box is seated. Beside the box is the position *mirror*, the start and endpoint of each task. The table is placed in the middle of the wall to the left of the mirror. Around the table, the positions *table centre*, *table right* and *table left* are located. At the opposite wall of the mirror, the scale is placed. In front of the scale is the position *scale*.

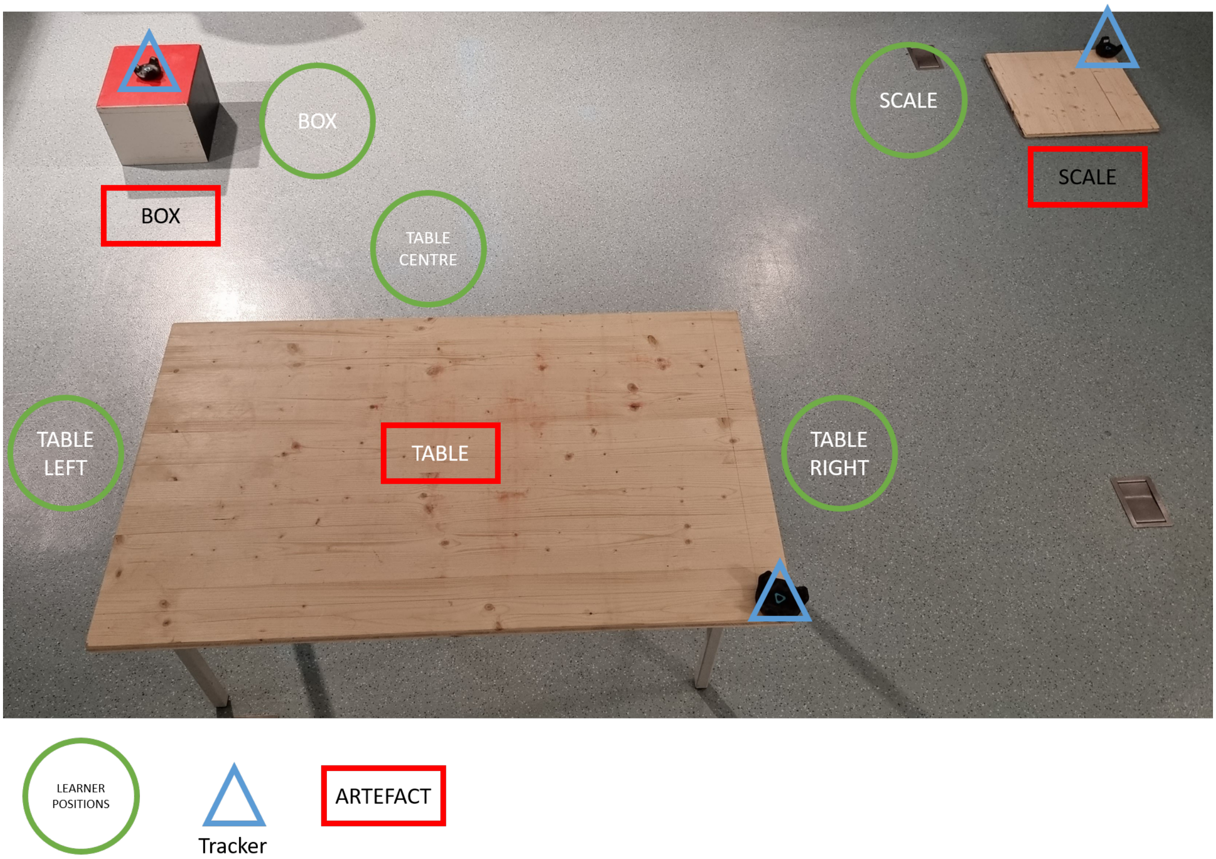


Figure 4.6.: Learner positions. Green circles indicate positions of the learner mentioned in the task description in table 3.3. Red suqares mark artefacts, blue triangles mark Vive Trackers

4.4. Guidance Visualisation

The next task is to add the GV to E(x|g)o, which the learner will mimic. The GV is an avatar like the learner's avatar, with the difference that the motion of the GV is driven by the pre-recorded tasks 1-3. The recording of the tasks was also performed with E(x|g)o. To use E(x|g)o as a recorder, a copy of all trackers and the HMD is created. This recorder-copy is packed in one GameObject with the trackers as children. The parent GameObject is recorded during the performance of the movements. For the GV, a similar GameObject as the recorded GameObject is created and serves as Input for FinalIK. In this section, the main points of the process are described: the recording of the movements and the resizing of the GV to the size of the learner.

The movements in the task have to be performed ergonomically. The measures to evaluate ergonomic movements are the RM. To serve as a strong baseline, a professional for ergonomic movements should record the movements. Because of the COVID-19 pandemic, all attempts to record the movements by a professional failed. The whole laboratory was transported to a private facility, but because of temporal issues, the recording with the professional could not take place. Then the laboratory was transported to another private facility. Unfortunately, the room in which the laboratory was set up was not suitable for the recordings. The recorded movements by the professional had to be abandoned because of insufficient tracking coverage causing jitter. The laboratory was transported back to the university. Eventually, I was trained by a physiologist and recorded the movements by myself. The final recordings were examined by the physiologist. Overall, the movements were rated by the physiologist as "by and large correct". The back is not always straight or at the correct angle. In task 2 during a *push* and in task 3 during a *pull*, the feet are misplaced.

With the recording of the tasks at hand, the GV can be animated. For the ego-centric VP it is inevitable that GV and learner have the exact same size. Otherwise, the learner cannot perceive the GV correctly. Furthermore, the table, box, and scale must not resize. The solution is to record two sets of object synchronously. The first set contains the objects that have to be resized, namely the GV, the second set contains objects that must not be resized, namely the box, table and scale. The recordings were synchronised by a script, the playback of the animations in E(x|g)o, too. The resizing of the GV takes place in three steps, compare figure 4.7. First, the learner and the GV are calibrated. Then the height difference Δy is measured between the learner and GV. In the last step, FinalIK is removed from the GV, the animated GameObject containing all trackers is removed, then FinalIK reattached and calibrated with the resized GameObject containing all trackers. After this process, learner and the GV have the same size.

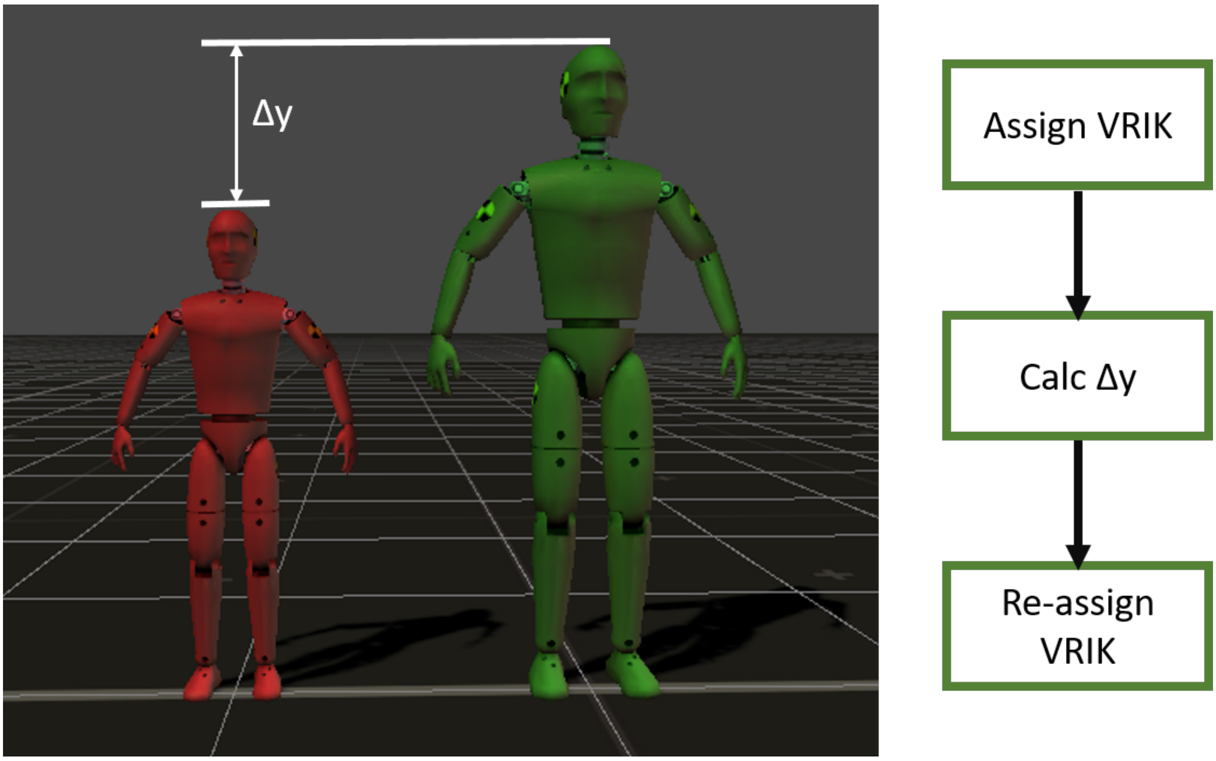


Figure 4.7.: Schematic description of the process to resizing the GV to the learner's height.

4.5. Visual Perspectives

The next element to add to E(x|g)o are the VPs the experiment will compare, namely the ego-centric VP, the exo-centric VP and the ego- & exo-centric VP. The implementation of the VP is partly informed by Chua et al. [21]. Chua et al. used full-body high realism degree avatar representations as GVs teaching stationary Tai Chi. In the ego-centric VP, the GV is placed inside the learner's avatar. In the exo-centric VP Chua et al. decided to placed the virtual copy of the learner and the GV side by side.

The ego-centric VP requires, besides the learner, one ego-centric GV. The learner needs to stay inside the GV. This is achieved by the *speed mechanic*, compare section 3.1.1. The learner's distance to the GV is calculated with the help of the tracker at the hip of the learner and the recorded tracker at the hip of the GV. The positions of the trackers at the hip are projected to the floor. The projection to the floor is necessary because the *speed mechanic* would apply if the GV bends the knees during *lift* and *lower*: if the GV bends the knees and the learner does not, the distance will increase between the two trackers in the y-component. This restricts the learner's ability to perform an error: if the learner does not bend the knees during *lift* and *lower*, the GV would stop and remind the learner to bend the knees. To investigate if the learner could percept to bend the knees, the learner must be allowed to make this error. This is why the *speed mechanic* relies on the distance between the two projected points on the floor.
Additionally, the distance used for the *speed mechanic* finds application in another functionality. In the ego-centric VP, the learner is located inside the teacher. This means that the learner's viewport is inside the head of the GV and let the learner see the inside of the GV head. This leads to distraction due to the partly rendered inner head. The solution is to remove the head rendering if the distance is below ETD_{max} and reinitiate the rendering above ETD_{max}. The rendering is removed by replacing the materials array of the head with a material array that contains only invisible materials. Figure 4.9 visualises the ego-centric VP by showing the bird's-eye view (top) and the view from the first-person perspective (bottom).

In the exo-centric VP, four exo-centric GVs are located around the learner. The positions of the exo-centric GV were determined after the task was recorded. The difficulty is to determine proper positions of the exo-centric GVs. First, at any point in time during every task's performance, the learner must be able to see a GV by only turning the head. Secondly, the GV and the learner should not move through a table or scale of another GV. The solution to the first part is informed by Chua et al. [21]. Chua et al. chose four representations that are in front, behind, left and right of the central learner. The second part proved to be impossible if the exo-centric representations should be near enough to be observed by the learner. A GV crossing through an artefact of another GV happens rarely and only during the subtask *carry*, but is a limitation of E(x|g)o. The GV needs to be shifted too far away from the learner not to cross other GVs artefacts. In a distance in which no crossing occurs, the movements are barely visible to be mimicked correctly. The exo-centric representations were then placed at a distance that allows being observed by the learner, and the learner can access all positions without standing in a digital artefact of another GV. The exo-centric GVs are positioned as follows. Standing at table centre and looking in the direction of the table: the GV to the left is shifted by two meters to the left, the GV to the right, two meters to the right. The GV in front is shifted by 1.5 meters to the front. The GV in the back is shifted 3 meters to the back. Figure 4.8 shows the positions of the exo-centric VP. Additionally, to the exo-centric GVs, a virtual copy of the student needs to be rendered. The same values shift the virtual copy of the learner. Figure 4.10 visualises the exo-centric VP by showing the bird's-eye view (top) and the view from the first-person perspective (bottom).

In the last VP, the ego- & exo-centric VP, the learner has an ego-centric VP and the exo-centric GV with the corresponding virtual copies of the learner. The implementation of the ego- & exo-centric VP is the union of the implementations of the ego-centric VP and exo-centric VP. Figure 4.11 visualises the ego- & exo-centric VP by showing the bird's-eye view (top) and the view from the first-person perspective (bottom).

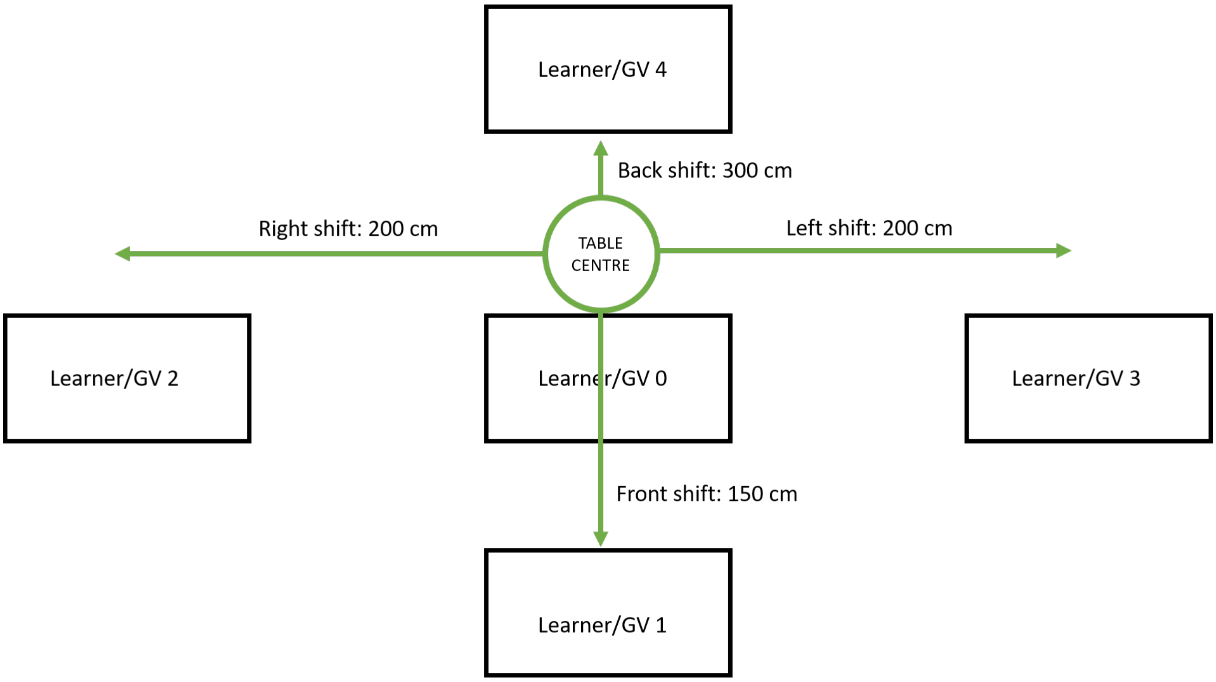


Figure 4.8.: Position of guidance visualisations.

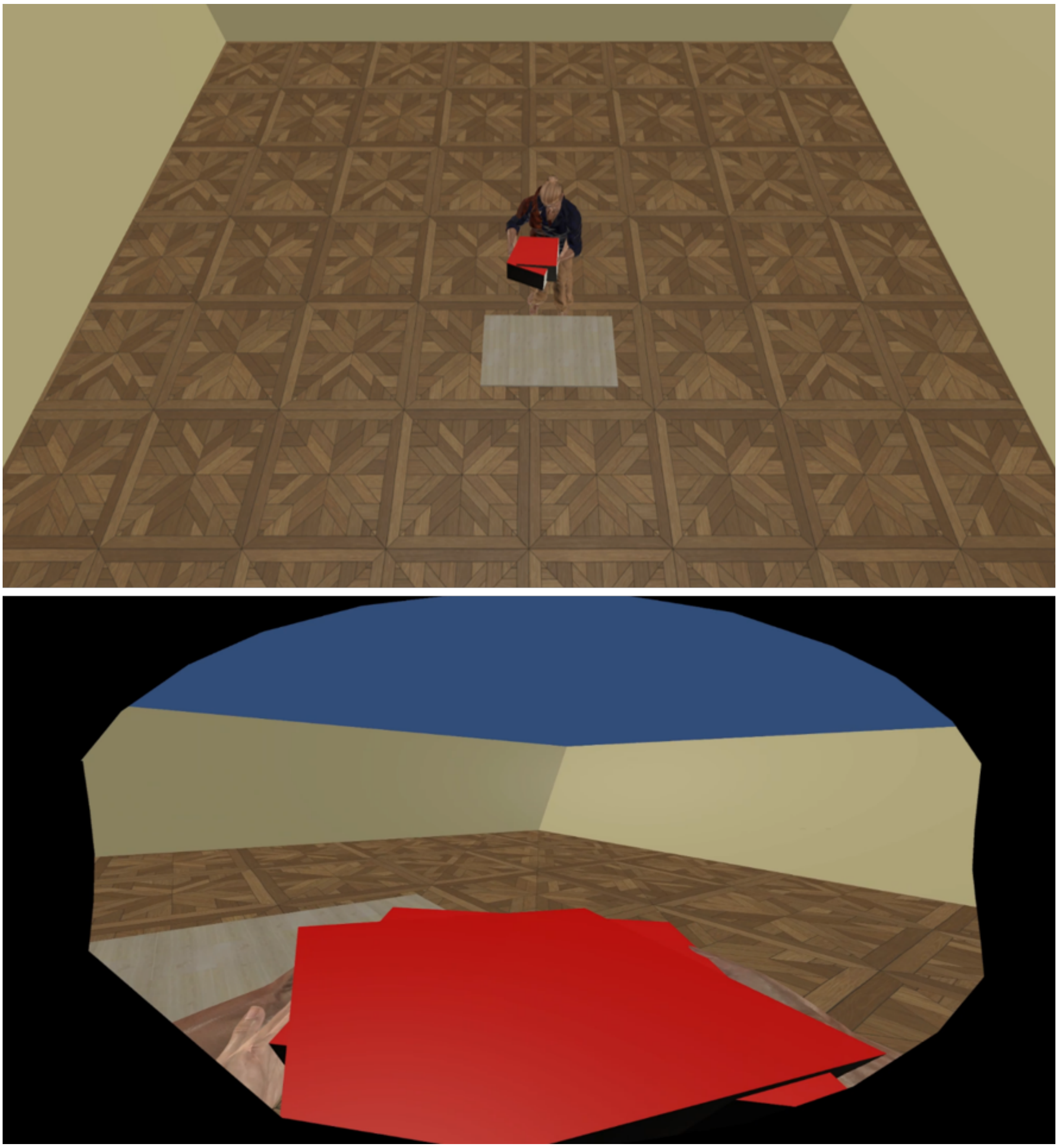


Figure 4.9.: Ego-centric VP from the bird's-eye view (top) and ego perspective (bottom) on the exact same scene. The GV (woman in red shirt) in located inside the learner (man in blue shirt).

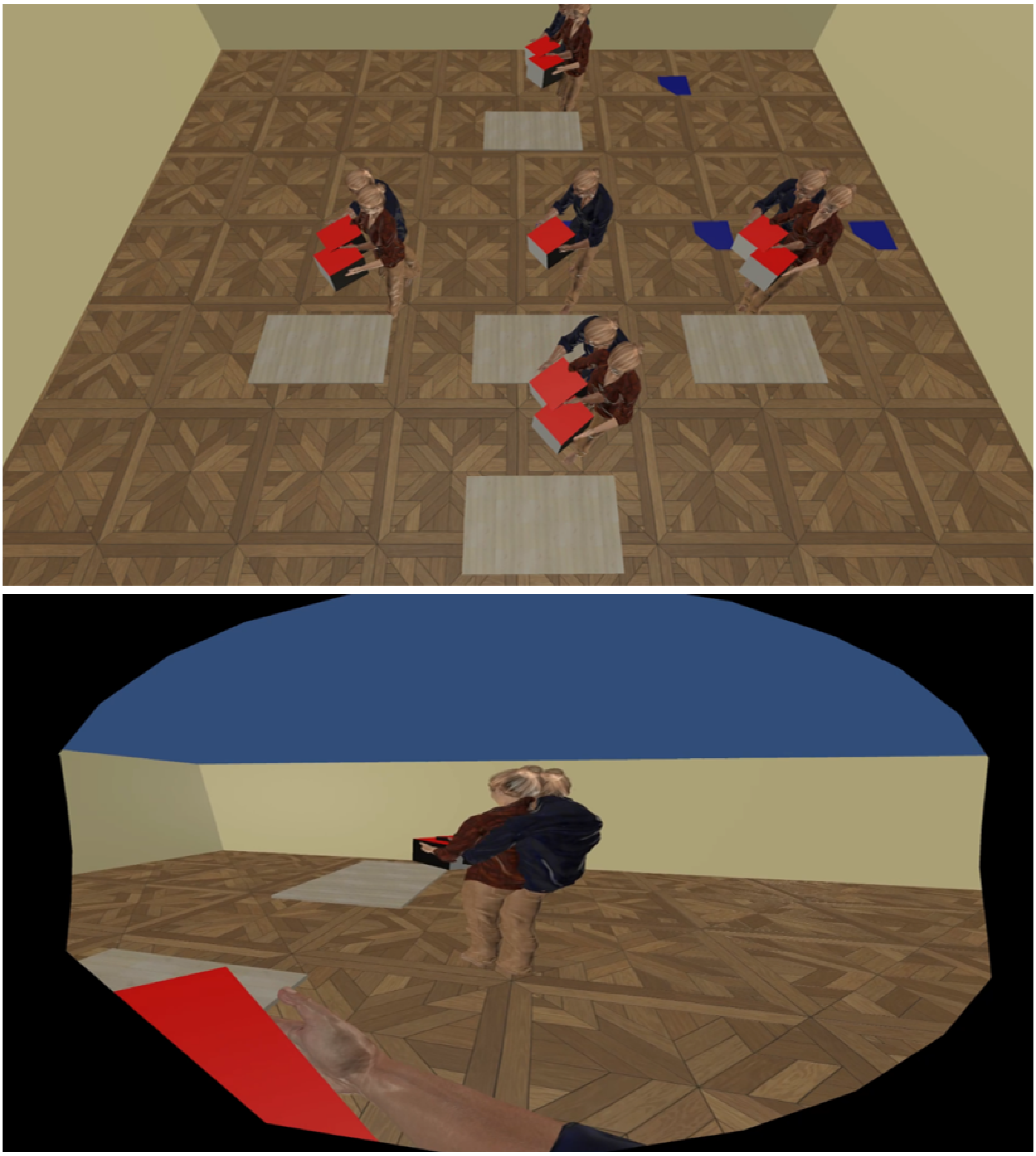


Figure 4.10.: Exo-centric VP from the bird's-eye view (top) and ego perspective (bottom) on the exact same scene. The GV's (woman in red shirt) are located around the learner (man in blue shirt). Additionally, a virtual copy of the learner is located inside the exo-centric GVs.

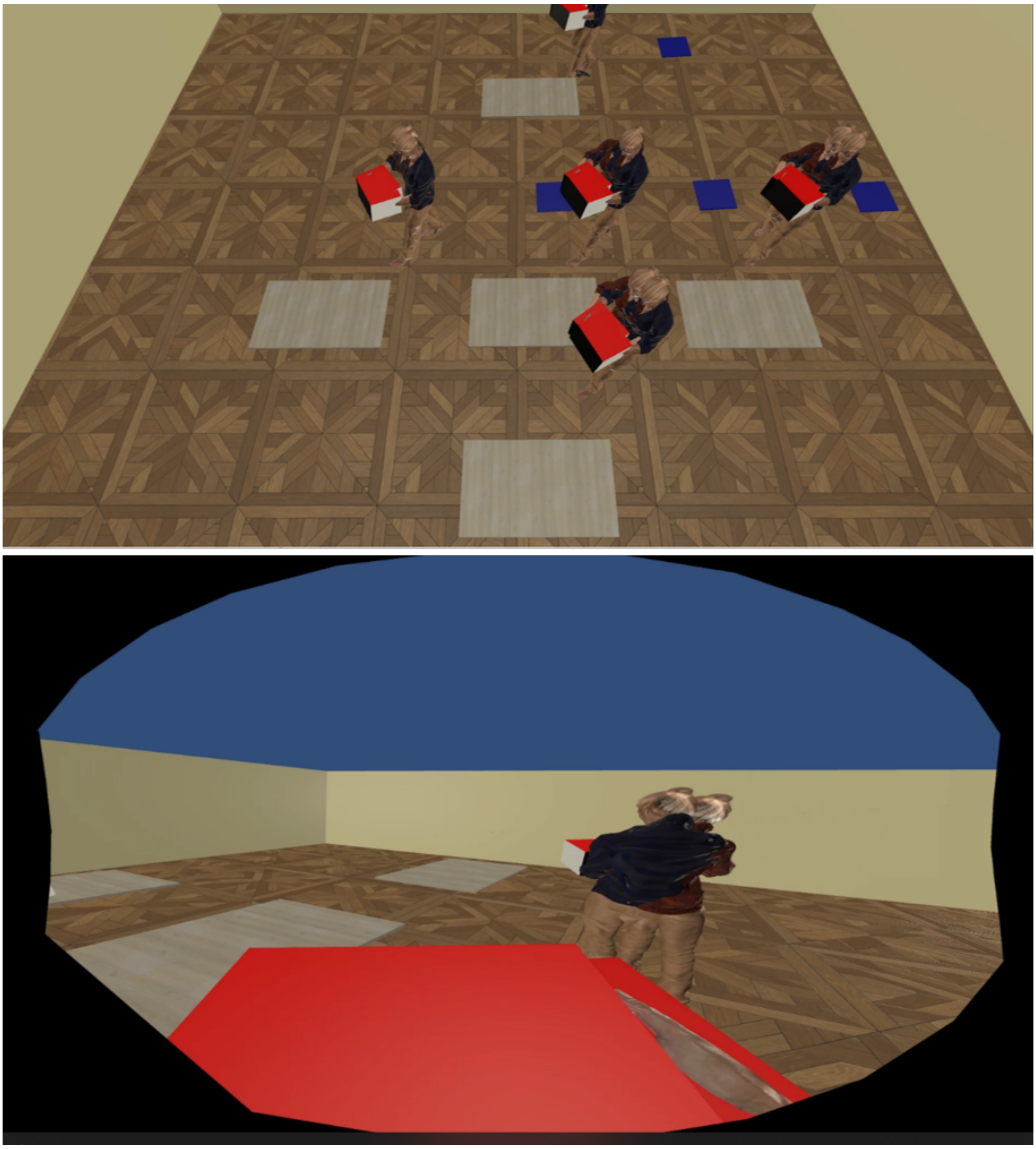
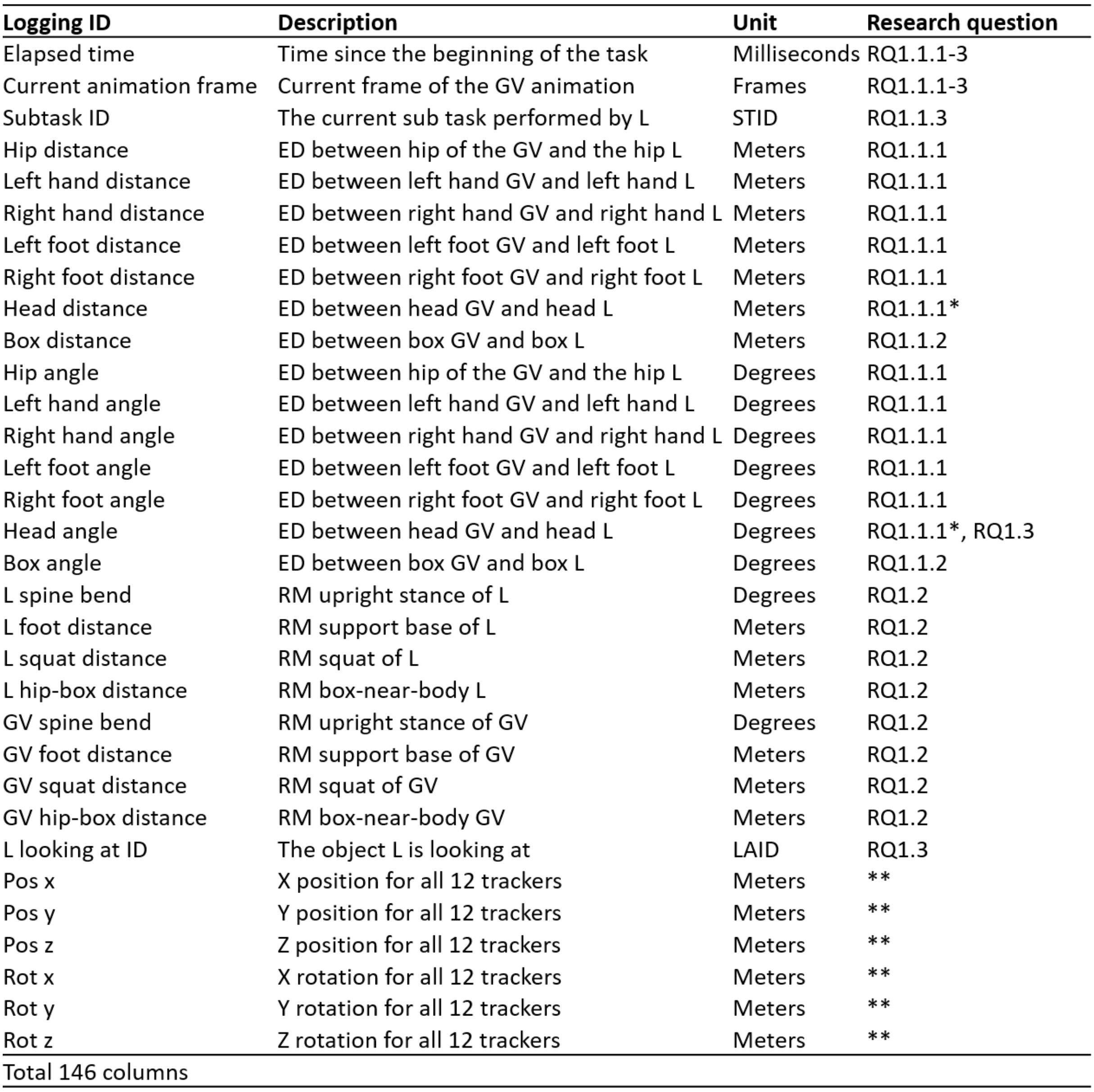


Figure 4.11.: Ego- & exo-centric VP from the bird's-eye view (top) and ego perspective (bottom) on the exact same scene. The GV's (woman in red shirt) are located around the learner (man in blue shirt) as well as inside. Additionally, a virtual copy of the learner is located inside the exo-centric GVs.

4.6. Quantitative Data Aquisition



Logging ID	Description	Unit	Research question
Elapsed time	Time since the beginning of the task	Milliseconds	RQ1.1.1-3
Current animation frame	Current frame of the GV animation	Frames	RQ1.1.1-3
Subtask ID	The current sub task performed by L	STID	RQ1.1.3
Hip distance	ED between hip of the GV and the hip L	Meters	RQ1.1.1
Left hand distance	ED between left hand GV and left hand L	Meters	RQ1.1.1
Right hand distance	ED between right hand GV and right hand L	Meters	RQ1.1.1
Left foot distance	ED between left foot GV and left foot L	Meters	RQ1.1.1
Right foot distance	ED between right foot GV and right foot L	Meters	RQ1.1.1
Head distance	ED between head GV and head L	Meters	RQ1.1.1*
Box distance	ED between box GV and box L	Meters	RQ1.1.2
Hip angle	ED between hip of the GV and the hip L	Degrees	RQ1.1.1
Left hand angle	ED between left hand GV and left hand L	Degrees	RQ1.1.1
Right hand angle	ED between right hand GV and right hand L	Degrees	RQ1.1.1
Left foot angle	ED between left foot GV and left foot L	Degrees	RQ1.1.1
Right foot angle	ED between right foot GV and right foot L	Degrees	RQ1.1.1
Head angle	ED between head GV and head L	Degrees	RQ1.1.1*, RQ1.3
Box angle	ED between box GV and box L	Degrees	RQ1.1.2
L spine bend	RM upright stance of L	Degrees	RQ1.2
L foot distance	RM support base of L	Meters	RQ1.2
L squat distance	RM squat of L	Meters	RQ1.2
L hip-box distance	RM box-near-body L	Meters	RQ1.2
GV spine bend	RM upright stance of GV	Degrees	RQ1.2
GV foot distance	RM support base of GV	Meters	RQ1.2
GV squat distance	RM squat of GV	Meters	RQ1.2
GV hip-box distance	RM box-near-body GV	Meters	RQ1.2
L looking at ID	The object L is looking at	LAID	RQ1.3
Pos x	X position for all 12 trackers	Meters	**
Pos y	Y position for all 12 trackers	Meters	**
Pos z	Z position for all 12 trackers	Meters	**
Rot x	X rotation for all 12 trackers	Meters	**
Rot y	Y rotation for all 12 trackers	Meters	**
Rot z	Z rotation for all 12 trackers	Meters	**
Total 146 columns			

Table 4.1.: Detailed overview of logs produced by E(x|g)o per frame. L: learner, GV guidance visualisation, ED: euclidean distance. *head position and rotation is biased in exo-centric conditions because of multiple GV the L can focus on. **All trackers are logged for backup reasons: after the experiment is conducted, a measurement can become interesting that was not of importance before. With these values, any measurement can be calculated post-experiment.

Section 3.3 defined the measures that are necessary to answer the research questions. E(x|g)o must be capable of assessing all measures. This section explains how E(x|g)o assesses the measures. An overview of all measures is listed in table 4.1. Table 4.1 lists the logging ID, a description of what the measurement is measuring, the unit in which the measurement is measured and for which research question the measurement is assessed. Quantitative data acquisition can be divided into several classes: (i) accuracy measurements (1-5), (ii) ergonomic measurements (6), (iii) focus measurement (7) and (iv) time measurements (9). In the following, (i)-(iv) are discussed.

4.6.1. (i) Accuracy Measurements

The accuracy measurements assess the discrepancy between the movements of the learner and the movements of the GV. Accuracy measurements are subdivided into distance-based measures and angle-based measures. Distance-based measures rely on the Euclidean distance between the learner's body parts and the body parts of the GV. The reference point for the body part is the tracker, which is attached to the body part. The body parts are: hip, left hand, right hand, left foot, right foot and head. The distance between the learner's box and the box of the teacher is an accuracy measurement, too. Similar to the body parts, the distance between the two boxes is the Euclidean distance between the tracker of the box and the recorded tracker of the GV. Please note, the trackers are not visible to the learner during the experiment.
Angle-based accuracy measurements assess the discrepancy in orientation between the body parts and box of the learner and the GV. The angles are measured in degrees. The calculation of the angle-based measurements complies with the calculations of the distance-based measurements. This means the angle between the corresponding trackers is measured. To summarise: distance-based measurements assess the positioning's error, angle-based measurements assess the error in orientation. Similar accuracy measurements are used by [2, 3, 10, 20, 21, 23].

4.6.2. (ii) Ergonomic Measurements

Figure 4.12.: Calculation of the risk metrics. a) *upright stance* — angle between the upright vector and the vector from the L5 tracker to the T8 tracker, b) *squat* — Euclidean distance between L5 tracker and floor, c) *support base* — Euclidean distance between the feet tracker, d) *box-near-body* — distance between L5 tracker and box tracker.

The ergonomic measurements are the four injury risk metrics: *support base*, *squat*, *upright stance*, and *box-near-body*.
support base is the distance between the feet, compare figure 4.12 c). For *push* and *pull*, *lift* and *lower*,

turn and *fold*, *pick* and *place* each a window in which the distance should be located is defined (see next of paragraph). The percentage of time the learner is inside the desired window is the outcome of the measurement. For a better understanding, imagine the following exemplenary scenario: the window during *push* for *support base* is 20cm - 30cm. During the performance of *push*, the learner's feet distance was inside the window for 90 seconds. The whole performance of *push* lasted 100 seconds. The RM *support base* yields in a score of 90%.

The measurement for *squat* is the distance between the hip and floor, compare figure 4.12 b). It indicates if the learner bent the knees correctly and is applied in the subtasks *lift* and *lower*. A window is defined for *squat*, too. Calculations of the RM score of *squat* complies with the *support base*.

Upright stance is the measurement of the spine bend, compare figure 4.12 a), which should be in a specific window, too. For *upright stance*, an additional tracker is applied to the back of the learner at Vertebra thoracalis 8 (T8), which is around 20cm kranial of the lower hip tracker at Vertebra lumablis 5 (L5)[12], compare figure 4.2 c). The angle of spine bend is the angle between the upright vector, and the vector of the upper hip tracker[13]. *Upright stance* is applied for *push* and *pull*, *lift* and *lower*, *turn* and *fold*. The bend angle during *pick* and *place* depends on the box' position on the table and thereby varies. Because of this variation, a window cannot be defined for *pick* and *place*, and thus the RM *upright stance* is not applied to *pick* and *place*. Calculations of the RM score of *upright stance* complies with the calculations of the preceding RM.

Box-near-body is the Euclidean distance between the hip tracker and the box tracker, compare figure 4.12 d). It is applied for the subtask *carry*. The calculations comply with the preceding RMs. The limitation of *box-near-body* is that the measurement is influenced by the circumference of the learner's torso. A formative test of *box-near-body* was not possible due to the COVID-19 pandemic.

The definitions of windows for the RMs were planned to be done by a professional with reasonable knowledge about ergonomics. Unfortunately, it was not possible to invite the professional to the laboratory to define the windows for the RMs. This means for the pilot study that the RMs cannot be evaluated.

---

[12] Latin description: Dr. med. univ. Kilian Roth

[13] Implementation for the calculations of the spine bend angle informed by Tanveer Singh Mahendra.

4.6.3. (iii) Focus Measurement

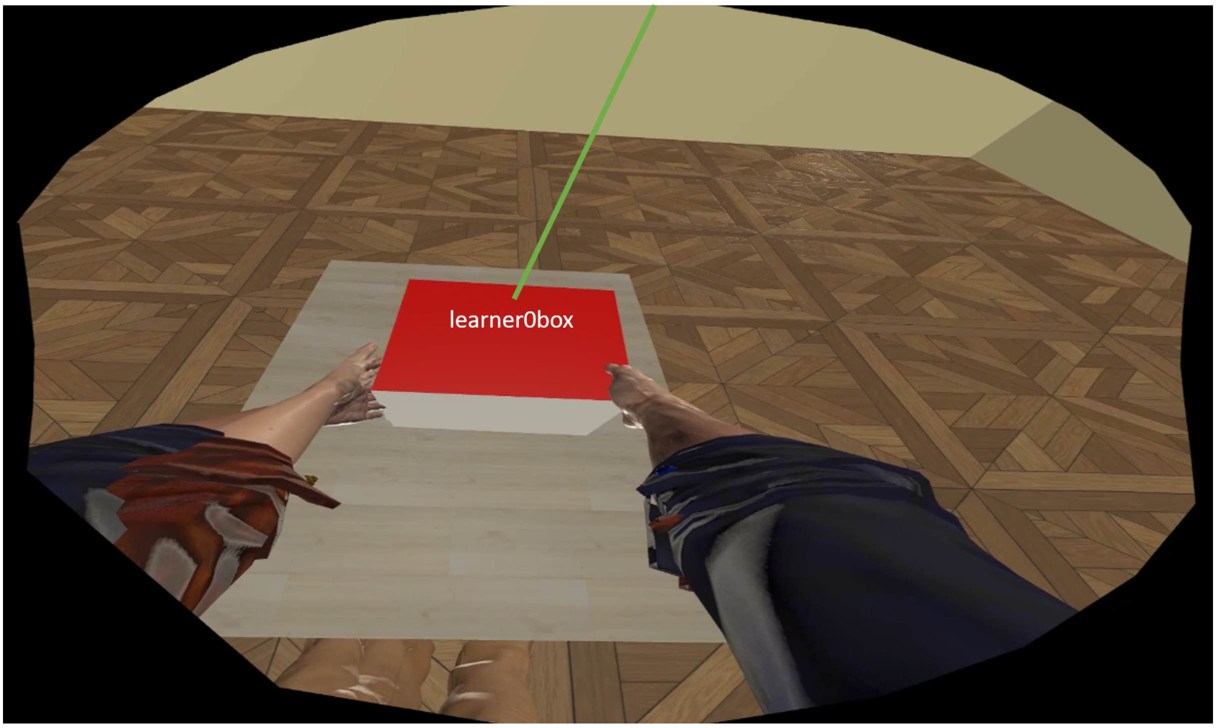


Figure 4.13.: Visual focus assessment. Ego perspective of the learner. The ray is depicted with a green line.

The virtual room the learner sees in E(x|g)o is filled with tables, boxes, GVs and virtual copies of the learner. To assess on what the learner is focussing during the tasks, every object was given a name. In every frame, raytracing is performed. The ray's origin is the HDM and expands straight forward. The name of the object first hit by the ray is written into the log file. The name is coded with the position 0-4 (see positions in figure 4.6) and an object identifier (box, table, scale, GV, learner), compare figure 4.13. A test revealed a systematic error of the ray pointing too high. To correct the discrepancy, the colliders of the objects were increased. The tables' and scales' collider height is increased by 20cm. The box' collider height was doubled. The learners' and GVs' avatar were wrapped into a capsule collider with a height 200cm of and a radius of 30cm. The values were determined by experimentation. To test the values, all subtasks were performed, and the object which was hit by the ray was displayed. The displayed name complied with the object in focus at nearly any point in time. Using an eye-tracker would increase the accuracy but was unavailable.

4.6.4. (iv) Time Measurements

The animation speed of the GV is determined by the distance between the learner and the GV (*speed mechanic*). A slower played GV animation results in a longer task completion time. For ease of un-

derstanding, please consider the following two definitions: the time the task lasts without the *speed mechanic* is defined as *task norm duration* (TND). The time the learner needs more than the TND to fulfil the task is defined as *over task norm duration* (OTND.)
OTND can draw conclusions about the learner's position in relation to the GV. The tasks differ in the amount of time to be performed:

- TND task 1: 172128ms
- TND task 2: 189040ms
- TND task 3: 176668ms

The OTND can be applied to specific subtasks, too. This measurement will mainly be used in the evaluation for triangulation.

Using the TCT for the evaluation of movements was previously done, for example by [2, 20, 29].

4.7. Qualitative Data Aquisition

The qualitative data assessment relies on one *questionnaire* after each *run* and a semi-structured *interview* after all three *runs*. The qualitative data assessment aims to determine the participant's impressions and opinions about the VPs. In the questionnaires, a different wording is applied to ease understanding. For example, the GV is called virtual teacher.

4.7.1. Questionnaire

The *questionnaire* (Appendix B.2.3) starts with a question about the subjective overall performance of the learner.

Q1: How accurate did your movements match with the virtual teacher?

A: Likert scale from one (very good) to 7 (very poor)

- Linked research questions: RQ1.1.1-3, RQ1.4
- Data triangulation for (1-4,6)

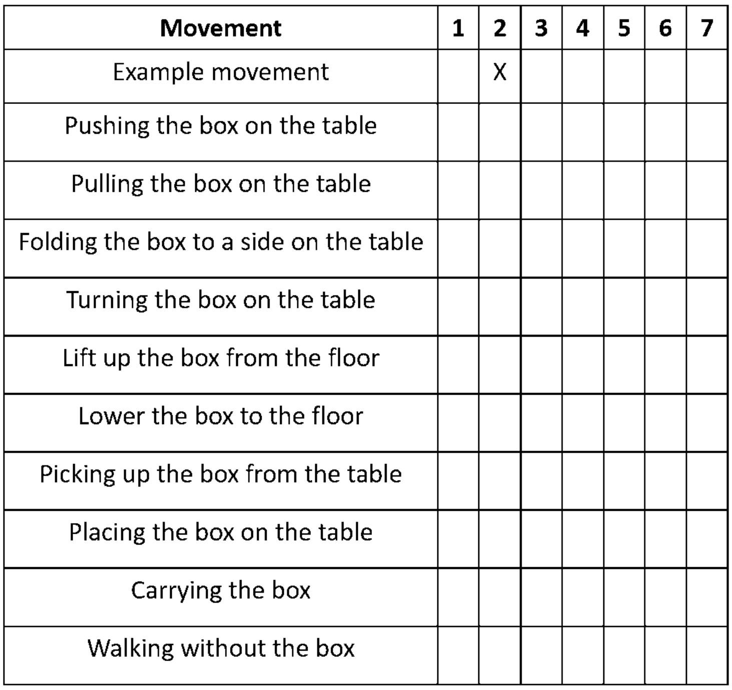
The answer to this question gives insights into how accurate the participant judges the performed movements. Furthermore, this question can be used to determine if the qualitative accuracy complies with the participants' subjective opinion.
The next question aims to assess the user's subjective performance for the subtasks. The participant is

asked to fill in a table. Each line represents a subtask. Each subtask can be rated on a Likert scale from 1 (very good) to 7 (very poor).

Q2: During the task, there were several smaller reoccurring movements, like pulling or lifting the box. Please rate these smaller movements to what extent you could follow the movements: 1 (very good) to 7 (very poor).

A: Likert scale from 1 (very good) to 7 (very poor) for each subtask.

- Linked research questions: RQ1.1.1-3, RQ1.4
- Triangulation for (1-4,5,6)



Movement	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Example movement		X					
Pushing the box on the table							
Pulling the box on the table							
Folding the box to a side on the table							
Turning the box on the table							
Lift up the box from the floor							
Lower the box to the floor							
Picking up the box from the table							
Placing the box on the table							
Carrying the box							
Walking without the box							

Figure 4.14.: Rating template for the subtasks.

Beside the subjective opition about the participants' performance, the answers of Q2 can be used to be compared with the qualitative data.
Q3 aims to assess the subjective accuracy of the participants body parts.

Q3: Please rate to what extend you think you could align your body parts with the teachers body parts: 1 (very good) to 7 (very poor).

A: Likert scale from 1 (very good) to 7 (very poor) for each body part.

- Linked resarch questions: RQ1.1.1-3, RQ1.4

- Triangulation for (1-4,5,6)

Body part	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Example body part		X					
Legs							
Arms							
Back							

Figure 4.15.: Rating template for the body parts.

Q3 assesses how good or bad the participant could perceive the body parts of the GV.
The last question is not handed to the participant. It serves as the basis for a subsequent semi-structured interview question. It gives the possibility to dig into extreme values of the questions answered before and into incidents that occurred during the performance of the task.

Q4: (As interview question) Did you have problems to follow the instructions?

  - E.g. because you could not see some body parts?
  - E.g. bad perception related to the perspective?
  - Go into extreme values of this questionnaire!
  - Address critical incidences!

A: Notes of participants statements.

4.7.2. Semi-Structured Interview

After all three *runs*, the participant is interviewed with the help of the semi-structured interview guideline (Appendix B.2.4). The guideline contains seven main questions, some with additional hints to dig deeper or to lower the participant's entry threshold to start reporting.

Q5: You saw three visual perspectives: ego-centric, exo-centric and the combination. What do you think about these perspectives?

  - entry question, encourage participant to talk frank, address interesting statements.

Linked research question: RQ1.4

Q6: Prioritise the perspectives by how accurate you could follow the movements. (1 best to 3 worst)

  - Why did you prioritize this way?

  Linked research question: RQ1.4

Q7: Imagine you want to learn a movement in VR. Which perspective would you use for that?

  - Or would you use a totally different one?

  Linked research question: RQ1.4

Q8: Which of the three perspectives was the easiest to understand?

  - Was there a perspective that confused you?
    - What do you think caused the confusion?
  - Was there a perspective you did not understand right away?

  Linked research question: RQ1.4

Q9: What do you think are the advantages and disadvantages of the perspectives?

  Linked research question: RQ1.4

Q10: Could you see some body parts better or worse in the perspectives?

  - What about your legs, arms, back?
  - Could you detect that during *lift* and *lower* you should squat?
  - Could you detect that you should step back during *push* and *pull*?

  Linked research question: RQ1.4

Q11: Did you miss a feature?

  - Dig for improvements for E(x|g)o or experiment design.

Q12: (Space to ask for critical incidences if any occurred.)

4.8. Experiment Procedure

E(x|g)o is now complete, and the elements of the experiment design can be assembled with the technological elements of E(x|g)o to form the final experiment procedure.
As soon as the participant enters the room, the participant receives a warm welcome to feel comfortable. The process starts with a welcome letter[14] (appendix C), followed by the informed consent and a demographic questionnaire. In the meantime, E(x|g)o is set up by choosing the condition, set the gender of the participant as well as the log is configured with the participant ID and task ID. After the demographic questionnaire, a spoken explanation about what is about to happen is given. Then the trackers are attached to the participant. The calibration process is explained. An explanation of the perspective is provided. Then, the first *run* is started. E(x|g)o gets started, the cameras and screen recording are set up, and the participant puts on the HMD. The participant is invited to calibrate. For calibration, the key C is pressed at the PC. To identify the camera recordings, a sign is held into the cameras. The task is started with the key S. During the *run*, the study conductor pays attention to the cable of the HMD to keep the participant from stumbling over it. Furthermore, the participant is observed. After the *run* ended, the HMD is removed. The participant fills in the *questionnaire. Run* two and three are conducted likewise. After all three *runs*, the trackers are removed. The participant is interviewed. The payment is given, and the receipt signed. At the very end, the participant is thanked and said goodby. If it appears, doorstep talk is appreciated.

4.9. Limitations

E(x|g)o and the experiment is designed for a task that includes the handling of physical load. If the results of the experiment can be applied to movements without a physical load is questionable. Additionally, it cannot be assumed that the results of the experiment can be transferred for other physical loads that are significantly different in shape and weight. Furthermore, the exo-centric GV sometimes walk through artefacts (table, scale) of other GVs, which can confuse the experiment participant. The movements are not recorded by a professional, errors in ergonomic movements are possible. Furthermore, the subtasks have a specific magnitude. They can not stand for the same subtask with different magnitude. This means, for example, for *lift*, the application of the outcome of the experiment for lifting up a box above the head is limited. Lastly, only a small number of participants participated in the formative tests to evaluate partial aspects of E(x|g)o. Especially, the hip-box distance is not tested because multiple persons with different physique would be necessary. The artefact contribution of this master's thesis, locomotion guidance in the ego-centric VP, is limited by being evaluated by one participant.

---

[14] Welcome letter, informed consent and demographic questionnaire partly informed by Daniel Schweitzer.

5. Experiment Evaluation

The experiment presented in the previous chapter is evaluated with the help of a pilot study. For practising the conduction of the pilot study, one participant was invited before the actual pilot study. The pilot study was conducted with three male participants between 25 and 29 years from the Computer Science department. All participants are experienced with VR devices, have rudimentary knowledge about the ergonomics of movements and rarely carry out tasks that include the handling of physical load. All of them are near-sighted, one of them wore glasses beneath the HMD. One participant previously used digital guidance to conduct sports ergonomically. The purpose of the pilot study was to identify flaws in the experiment procedure (section 5.1), E(x|g)o — including the assessment of the quantitative data (section 5.2) — and in the assessment of the qualitative data (section 5.3). Furthermore, the eligibility of the acclimatisation method is discussed in section 5.4. Finally, a glimpse on the data produced by the pilot study is provided in section 5.5.

5.1. Procedure

In the beginning, the participant receives a welcome letter. Two participants stated that the information in the welcome letter (Appendix B.2.1) about how VR headsets work is not necessary. The welcome letter was then shortened by the removal of that passage. The informed consent is signed. The following demographic questionnaire allows putting the study's data into context. During the review of the produced study data, no further questions occurred. The demographic questionnaire needs no further improvements. While the participant reads the welcome letter and fills in the demographic questionnaire, the system is set up. The pilot study showed that the time to set up the system and reading the welcome letter, and filling in the demographic questionnaire, are corresponding.

Afterwards, the participant is told that he/she will be equipped with the trackers and where the trackers will be attached. Afterwards, the trackers are attached to the participant. To respect the privacy of the participant, the trackers are handed to the participant and instructions are given on how to attach them. The pilot study showed that the participants had problems following the instructions correctly. For the actual study, the participants should be asked if it is okay that the trackers are attached with the physical help of the study conductor.

Next, the participants received information about what to expect in the VR. The instructions contained information about the GV ("You will see one/multiple teachers.") and the task ("Please follow the instructions of the teachers as exactly as possible."). Furthermore, the participant was asked to pay attention to the ergonomics of the movements. Explanations about the *speed mechanic* were provided, too ("The teacher will wait for you if you are too far away from the teacher. If that is happening, correct the placements of your feet, and it will go on."). No participant had difficulties understanding the instructions.

Additionally, the participant was handed the box to get used to it before seeing its digital pendant in VR. Finally, the calibration of the system was explained ("Please look into the mirror which you will see there (study conductor pointing) and extend your arms like this (study conductor performing the T-pose)"). All participants understood how to calibrate easily. The introduction of the mirror as a calibration facilitator proved to be helpful and suitable.

Subsequently, the camera recordings were started, the participant put on the HMD, and the participant performed the first task. In this phase of the study, two errors occurred which needed adjustments to the process. In one case, the wrong task was chosen, which made participant 1 (PT1) perform task 1 (T1) two times in different perspectives. PT1 recognised that, too. Before starting the task, an additional checklist should be gone through to ensure the correct task and perspective is chosen by the study conductor. The second error regards the identification and synchronisation of the video recordings. At the ceiling, a GoPro records the scene from above. A second camera catches the scene from the side of the tracking volume. For identification, a sign was held into the view of both cameras. This was forgotten twice by the study conductor. As an improvement, the sign should be placed beforehand in the area both cameras cover.

After the participant performed the first task, the participant took off the HMD and is asked to fill in the *questionnaire*. The trackers stayed at the body of the participant. The pilot study showed that the tracker did not hinder the participants from sitting down and filling in the questionnaire. In the pilot study, a three-minute pause was conducted to allow the participant to recover. During that pause, the participant was asked about his/her wellbeing to check for VR induced motion sickness. All participants stated that they do not need a pause. The demographic questionnaire revealed that all participants were experienced with VR-system and they are used to wear VR HMDs. The pause will be maintained because a person with no prior exposition to VR could feel different.

Run two and three are conducted in the same way as *run* one. With all three *runs* done, the trackers are removed, and the semi-structured interview was conducted. Because the pilot participants were not paid, the pilot study ended here. The planned duration of the study was 75 minutes. All pilot studies took no longer than 55 minutes. However, a time buffer should be planned in case some participants need more time for the experiment. With an additional buffer of ten minutes, the planned study duration can be decreased to 65 minutes.

Additionally, for the study's evaluation, the participants were interviewed to get insights about the experiment's and system's flaws. The participants were asked if the explanations were sufficient and if there were any confusing elements or unclear questions in the questionnaires or documents. Finally, the participants' opinion about possible improvements of E(x|g)o and the study were asked. The results of those interviews informed section, too.

To conclude, vast parts of the planned study process proved to be suitable. Adjustments are made to the welcome letter, the trackers' attachment with the study conductor's help, an additional checklist to check the *run's* task and perspective is introduced, and the camera recording identification is improved by placing the sign into the recording area beforehand.

5.2. E(x|g)o and Quantitative Data Assessment

All hardware artefacts of E(x|g)o are suitable without objections. The study participants rated the box' size and weight as "okay" while still perceiving it as a physical load. The table's size is sufficient for all three tasks. At no time, the participants were in danger to collide with a physical artefact. The size of the scale is also sufficient. The box was always placed on the scale safely. The positions and itinerary between mirror, table and scale are without complaint. Regarding the hardware part of E(x|g)o, the pilot study revealed two insights, one related to the trackers and one related to the HMD.

The tracker at the hip is attached with a strap around the hip of the participant. While the subtasks *lift* and *lower*, the tracker is shifted upwards. The upwards shift affects the avatar's presentation and influences the accuracy measurements for the hip and the RM *squat* and *upright stance*. To prevent this, the student was asked to wear a belt. The tracker belt was then fixated to the participants' belt with a band of velcro. This includes touching the participants in the lower hip area from behind. To prevent participants from feeling uncomfortable during the whole study, the fixation of the two belts should be performed with a clip that the participant can attach themselves.

The second insight regarding the hardware of E(x|g)o is the cable of the HMD. During the study, the study conductor handled the cable not to influence the participant. In one case, the cable was plugged out during a the performance of the participant. E(x|g)o is designed for that case, and plugging in the cable again allows to continue with the *run*. However, in the actual study, this incident would lead to unusable data for all three *runs* of that participant, because meanwhile the cable is plugged in again, the GV will move forward, and the error will be high during this phase. The actual study will benefit from a wireless HMD.

One participant stated that he/she could not identify the ownership of the box right away. As soon as the own box is in the participant's hands, it is no problem to tell which is the GV's box and which is the learner's box. However, if both boxes are stationary, the participant could not detect which is the own box. The box' appearence should be changed to light transparency for a better distinction between the learner's box and the GV's box. This will also have an influence on the perception of the box during *lift*. During *lift*, the learner's box is occluded by the GV's box for a short time. For conformity, the avatar, table and scale of the GV's should also be rendered with light transparency.

The pilot study served as the last test before the actual study. An essential part of the pilot study is the review of the produced data. During the development, the measures could only be tested individually. The pilot study allowed for the first time to get a whole image. Fortunately, most of the logged data worked as intended. Only minor errors were detected. For *upright stance*, a last-minute edit caused an incorrect calculation. For EXO, in combination with task 2, an invisible error was detected: the props animation controller, which animates the GV's box, played the wrong task for the ego-centric GV. In EXO, the ego-centric GV is invisible but used to calculate the measurements. For EGO, the learner's avatar identification name (used for *looking at*, identifying what the learner is visually focussing on) is incorrect. Lastly, Unity3D natively uses comma as decimal separator. Most statistics programs natively use the point as decimal separator. A log file of one *run* contains around 2.5 Million decimal separators. Converting a log file is time-consuming, and therefore, the logging should be changed to use a point as decimal separator.

5.3. Qualitative Data Assessment

After each task, the participant fills in the *questionnaire* (Appendix B.2.3). The last question of the questionnaire is not handed to the participant but asked personally. The feedback of the pilot study's participant was positive, and the questions were clear and understandable. However, during the analysis of the quantitative data, some aspects became interesting to consider the opinion of the participants. Therefore, the *questionnaire* is extended by the following questions:

Q: How ergonomic do you think your movements have been?

A: Likert scale from 1 (very good) - 7 (very poor)

Linked research questions: RQ1.2

The self-assessment of the study participants' performance regarding the ergonomics of the movements can give insights into the participants' opinon. It is expected that the VPs with exo-centric GVs is rated higher because the posture of the GV can be observed from the side. For example, the back of the GV's bend angle is hard to see in the ego-centric VP. Furthermore, the self-assessment can be put into relation to the quantitative data and used for triangulation.

Q: On what did you focus most?

A: Two choices: box, teacher

Linked research questions: RQ1.3

The assessment of the subjective participants' focus can give insights about the importance of box and avatar for the participants. Presumably, a participant who rates the box over the avatar focussed mainly on the box's accuracy and vice versa. This subjective data can be put into relation to the quantitative data for triangulation. If the quantitative data complies with the qualitative data, the evaluation of the risk metrics can be split into two groups: those who focussed mainly on the box and those who focussed mainly on the avatar. It could be interesting if these two groups score differently regarding the risk metrics.

Q: Could you see a teacher at any point in time?

A: Two choices: yes, no

Linked research questions: none, evaluation of the positions of the GVs

The positions of the GVs could only be tested sparsely. This question aims to evaluate if the positions of the GVs were suitable. If the participant crosses "no" in the last question, this can be addressed during the last personally asked question.

Q: Please rate to what extend you think you could align your box with the teacher's box:

A: Likert scale from 1 (very good) - 7 (very poor)

Linked research questions: RQ1.1.2

> The subjective assessment of the box allows a comparison with the quantitative data and gives insights to the participants' opinion.

After all three *runs*, the participant is interviewed with the help of the semi-structured interview guideline (Appendix B.2.4). The guideline for the semi-structured interview proved to be suitable. However, during the interviews, one participant could not give a clear prioritisation of which perspective he/she would use if he/she had the choice. I gave in and did not insist on the prioritisation. This caused an inconsistency which bacame clear during the analysis of the qualitative data. In the actual study, every participant should provide a prioritisation.

5.4. Acclimatisation Method

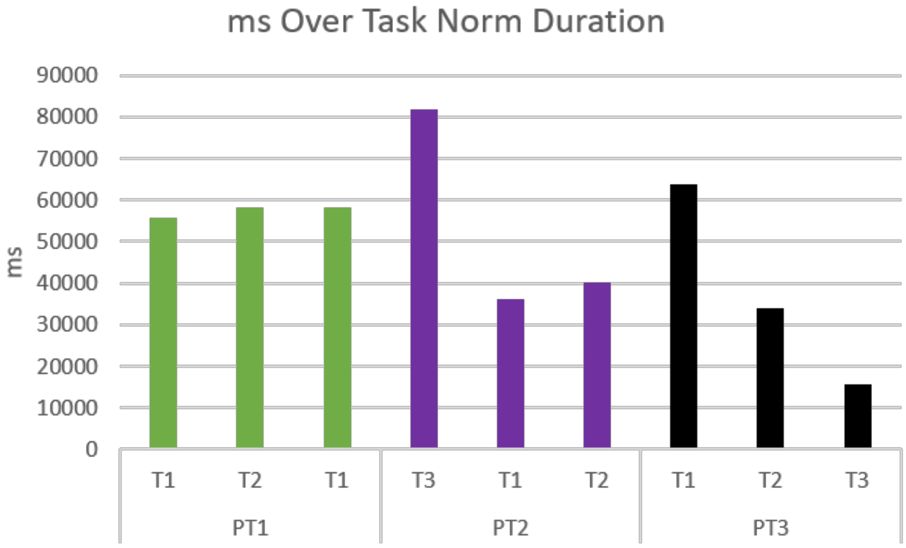


Figure 5.1.: Amount of miliseconds OTND per *run*. Prefix PT — participant, Prefix T — task.

The first task of each study is for acclimatisation, where the learner gets used to E(x|g)o. Thus, the data of the first task is excluded from the evaluation. It is assumed that the learning effect between *run* one and two is high, and between two and three, low learning effect occurs. To evaluate if this assumption holds, the task completion time (TCT) could give insights because of the *speed mechanic*. The *speed mechanic* regulates the GV animation speed based on the distance between learner and GV and is applied in all perspectives. The higher the learner-GV distance, the lower the speed. A learner who is often located near to the ideal point yields a lower TCT. Comparing the TCT in the pilot test could at least indicate if the learning effect between *run* two and three is low by showing similar *TCTs*. Figure 5.1 shows the amount of ms the participants needed more than than the TND to complete the task (OTND). The order of the *runs* is from left to right. Participant 1 (PT1) had a nearly equal OTND for all three tasks. Because of a mistake in choosing the correct task, PT1 faced T1 two times. However, in *run* three, the OTNT is slightly higher. PT2 shows the expected behaviour, having a high OTND in the first *run* and a nearly equal OTND for *run* two and three. PT3's OTND is strictly monotonically decreasing. If the OTNT behaviour of all the participants would be like PT2's behaviour, the choice of the acclimatisation method is correct. If the OTNT behaviour for all participants would look like PT3, the acclimatisation method would be incorrect. In this case, a separate condition specific acclimatisation before every *run* should be conducted. If the OTND behaviour for all participant would look like PT1, it could be discussed that no acclimatisation is necessary. Unfortunately, the data is ambiguous and does not allow an evaluation of the acclimatisation method. Thereby, using the first *run* for acclimatisation is maintained.

5.5. Data Analysis

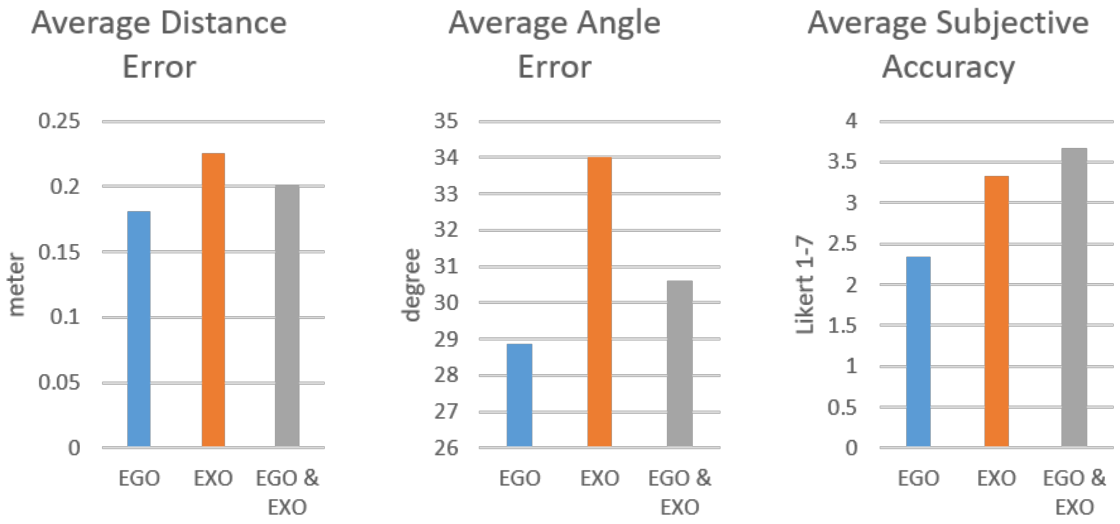


Figure 5.2.: Overall average error in distance (left) and angle (middle). Overall subjective accuracy (right).

Based on the pilot study, this section tries to give a first glimpse on the produced data. A pilot study serves to identify issues and faults in the system and experiment and prepare the final study conduction. The previous sections described found issues and faults and the proposed solution for them. Some issues and faults impacted the data, which led to the exclusion of corresponding data. Furthermore, as described in section 3.2.3, a full counterbalancing of tasks and conditions requires at least nine participants. Additionally, the data revealed, in some aspects, a high variation in both qualitative and quantitative data. Therefore, the depicted data is a rough estimation, and conclusions can not be drawn. The analysis is superficial, and a detailed analysis like significant verification is renounced. However, the first impression of a possible outcome can be given. All charts depicted in this section are similarly structured. The conditions in all charts have the same colour coding: EGO is depicted in blue, EXO is depicted in orange and the combination EGO & EXO is depicted in gray. For all charts (except for head

angle) holds: the lower the bar, the better in the corresponding context.
Figure 5.2 can be perceived as an abstract for this section by showing the overall average error in distance and angle between the learner and the GV per VP, as well as the subjective accuracy. The ego-centric VP outperformed both the exo-centric VP and the ego- & exo-centric VP in terms of accuracy, while the ego- & exo-centric VP mostly scored better than the exo-centric VP. The subjective accuracy is also highest in the ego-centric VP. In contrast to the objective accuracy, the participants rated their accuracy worst in the ego- & exo-centric VP.

5.5.1. Accuracy

Accuracy is clustered by distance and angle and applied for the body parts hands, feet, hip, head and box. Distance means the Euclidean distance between the learner's, e.g. hand and the GV's hand in meters, and describes the difference in position. Angle describes the difference in orientation and is measured in degrees. The overall average error per body is depicted in **??**.
Section 3.2.3 showed that some subtasks could be paired up, based on the similarity of the movements: *lift*/*lower*, *push*/*pull*, *turn*/*fold* and *pick*/*place*. Figure 5.4 shows the average error in distance and angle per subtask and confirms the pairing of the subtasks by showing a relation between the pairs. Thus, the pairs of subtasks will be analysed in combination. *Carry* and *walk* are analysed seperately. This section analysis the accuracy based on the body parts (section 5.5.1: Accuracy per Body Part) and the accuracy during the subtasks (section 5.5.1: Accuracy per Subtask) and compares the objective accuracy with the subjective accuracy.

Accuracy per Body Part

Figure 5.3.: Average error per body part. Left: distance error, right: angle error. Suffix D: distance, suffix A: angle. LH - left hand, RH - right hand, LF - left foot, RF - right foot. Bottom: subjective accuracy per body part, rated on a Likert scale 1 (best) - 7 (worst).

The **hip** error indicates to what extent the learner could determine the correct location. The data indicates that the determination of one's own position and rotation is more straightforward with an ego-centric GV.

Figure 5.3 shows a relation between the distance error of **hands**. This is expected since large parts have synchronised movements. For example, the hands touch the box simultaneously. The hands' error is lower in EGO than in EXO. Hands are directly visible in front of the learner, and the direct comparison to the ego-centric GV is a possible explanation for the higher accuracy in EGO. The subjective accuracy complies, being highest in EGO.

Surprisingly, **feet**'s error is lower in EGO, too. To see the feet, the learner must actively move the head, primarily if the box blocks the view on the feet. However, it seems easier to align the learner's feet with the GV feet in EGO. The subjective accuracy is lowes in EGO.

The **head** angle is not comparable with the other angle-based accuracy measures. The presence of multiple exo-centric GVs in the EXO forces the learner to look into different directions. The difference between EXO and EGO & EXO could point out that the learner focussed in EGO & EXO on both the exo-centric GVs and ego-centric GV. The angle-based accuracy for the hand and feet reveal no clear trend. More participants could provide a clearer view.

The **box** distance and angle error are lower in EGO and EGO & EXOthan in EXO. The presence of an ego-centric GV box increases the distance and angle accuracy for the box.

To summarise, the presence of an ego-centric GV increases the distance accuracy for all body parts. For the physical load, the presence of an ego-centric GV increases the angle accuracy, too. However, adding exo-centric GVs to an ego-centric GV increases the distance error. A possible reason why adding an ego-centric GV to exo-centric GVs increases the ddistance error could be, that the learner shares the focus with multiple GVs or that the presence of four exo-centric GVs overwhelms the learner. The study participants rated their overall accuracy highest in EGO and lowest in EGO & EXO, compare 5.2 (right). The low subjective accuracy in EGO & EXO would underpin the theory that the learners were overwhelmed. However, when the participants were asked about their subjective accuracy for the body parts, arms, legs and back, a different picture arises, compare 5.3 (bottom). The participants' opinion is differentiated, which causes a high standard deviation of around 1.5 on a Likert scale from 1-7. More participants could lead to a clearer view.

The described insights rely on the whole task containing all subtasks. Thus, the drawn deductions are only valid for the whole task. Potentially, the accuracy in specific subtasks could differ from the overall accuracy. In the following, the subtasks are analysed.

Accuracy per Subtask

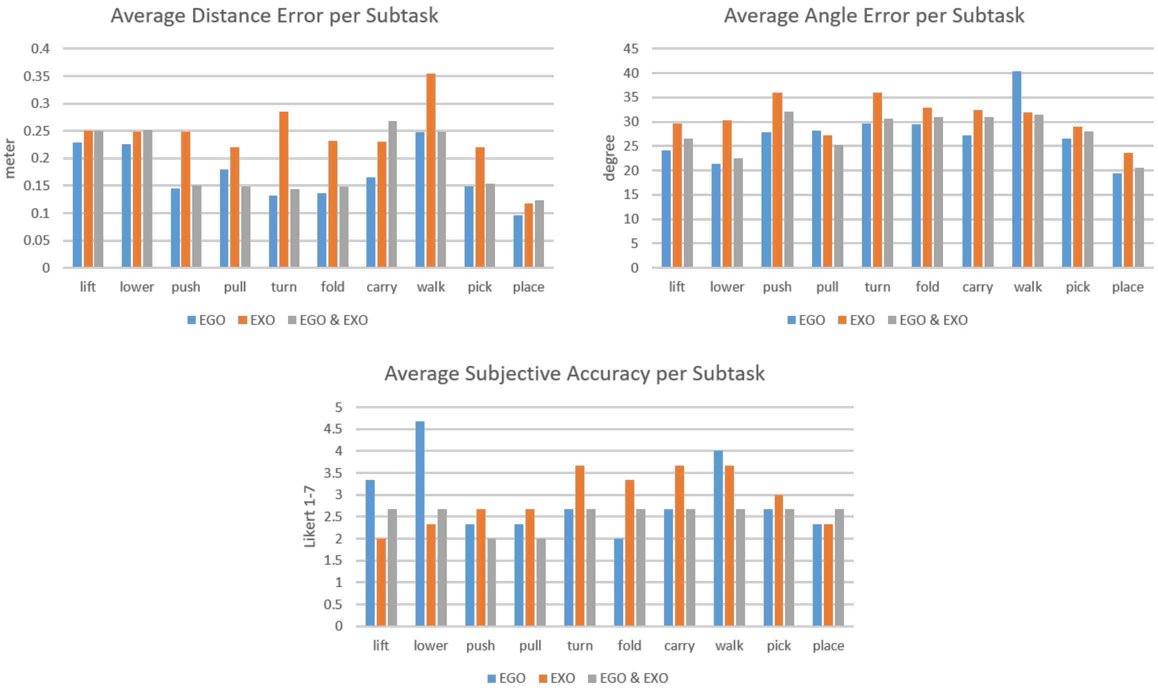


Figure 5.4.: Overall average error per subtask. Left: distance error, right: angle error. Bottom: subjective accuracy per subtask, rated on a Likert scale 1 (best) - 7 (worst).

Figure 5.4 depicts the overall distance and angle error for all subtasks, as well as the subjective accuracy for all. The next sections discuss the subtasks' accuracy in detail.

lift/lower

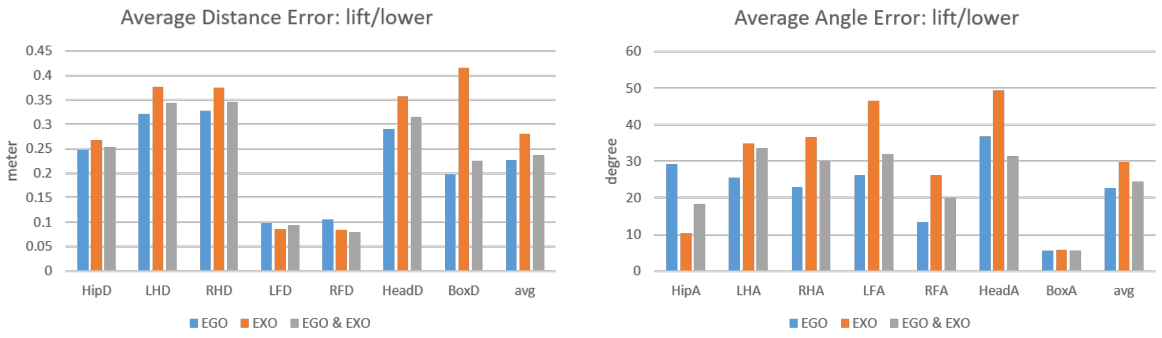


Figure 5.5.: Average error per body part for subtasks *lift/lower*. Left: distance error, right: angle error. Suffix D: distance, suffix A: angle. LH - left hand, RH - right hand, LF - left foot, RF - right foot.

Figure 5.5 shows that in EGO, the hip, hand and head accuracy is higher than in an EXO. The presence of exo-centric GV seems to have a positive influence on the feet's accuracy. The box' accuracy in EXO much is lower than in EGO and EXO. In the actual study, particular attention should be paid to the box during *lift* and *lower* to identify the cause of why EXO performed badly. In orientation, the box's error is low for all VPs. This is expected since the subtask does not include a change in orientation. The subjective accuracy is lowest in EGO, followed by EGO and EXO, compare figure 5.4 (bottom).

pick/place

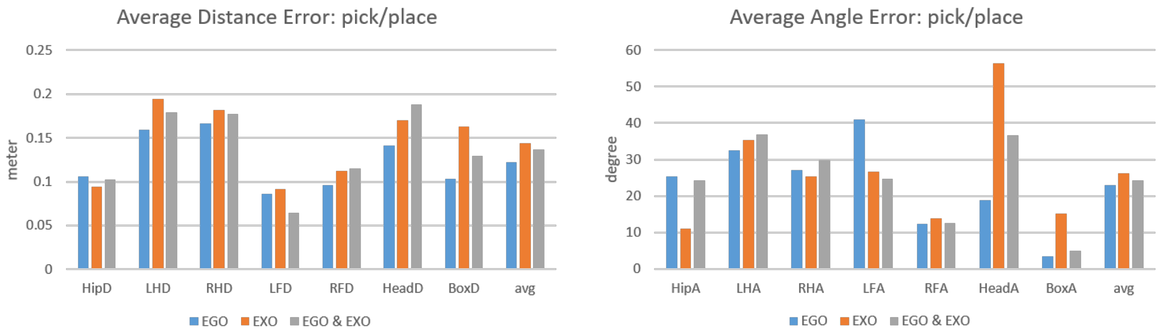


Figure 5.6.: Average error per body part for subtask *pick/place*. Left: distance error, right: angle error. Suffix D: distance, suffix A: angle. LH - left hand, RH - right hand, LF - left foot, RF - right foot.

Pick and *place* are *lift* and *lower* movements with a significant difference in magnitude. The accuracy of *pick* and *place* benefits from the presence of an ego-centric GV, compare figure 5.6. The distance and angle error of the box is lowest in EGO. The participants rated the accuracy equally for all conditions, compare figure 5.4 (bottom).

push/pull

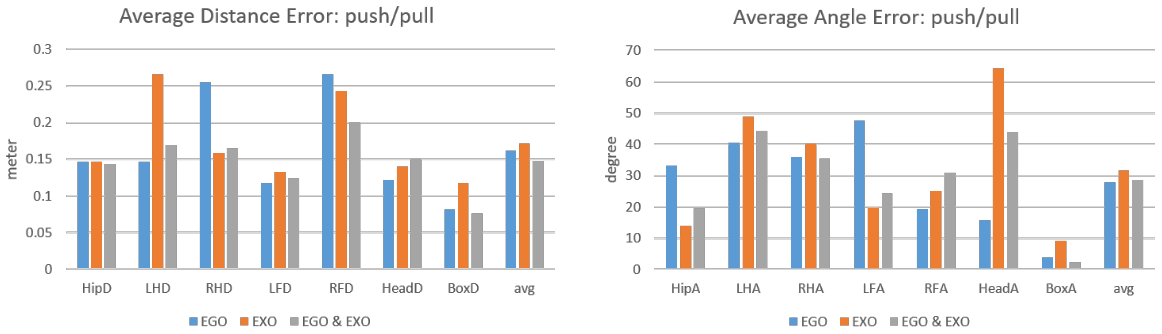


Figure 5.7.: Average error per body part for subtasks *push/pull*. Left: distance error, right: angle error. Suffix D: distance, suffix A: angle. LH - left hand, RH - right hand, LF - left foot, RF - right foot.

During *push* and *pull*, increased force is applied to the box. The physiologist suggested that one foot should be shifted to the back for *push* and *pull* because of the increased force application. The high difference in error between the left foot and right foot is based on different foot placement. Unfortunately, the participants realised the different foot placement not often. This is shown by the higher error of the right foot compared to the left foot, compare figure 5.7. One participant stated that he did not realise to shift one foot back in EGO in the interview. However, in EXO, he saw it and applied it then also for EGO & EXO. This statement harmonises with the quantitative data, which shows the lowest right foot accuracy in EGO. The left hand seems to have a high error in EXO, the right hand in EGO. The video revealed that the participants alternated the hand placement during *push* and *pull*. Based on the video observation, the high error could even out with more participants. The higher error of the head angle in EXO compared to EGO & EXO indicates that the participants shared the focus in EGO & EXO with the ego-centric GV and the exo-centric GV. The participants rated their movements more exact in EGO & EXO than in EGO and lowest in EXO, compare figure 5.4 (bottom).

turn/*fold*

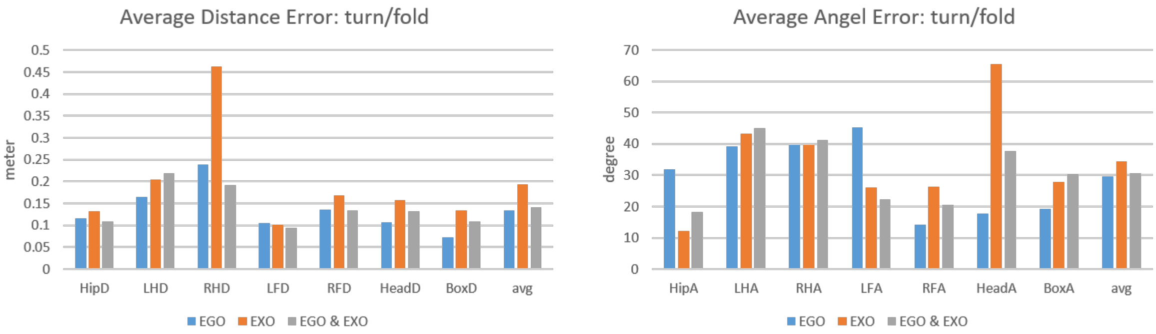


Figure 5.8.: Average error per body part for subtask *turn*/*fold*. Left: distance error, right: angle error. Suffix D: distance, suffix A: angle. LH - left hand, RH - right hand, LF - left foot, RF - right foot.

Most of the movements during *turn* and *fold* happens on the table. Thus the main focus is on the hands. The high error in EXO is noticeable, compare figure 5.8. The consultation of the video recordings revealed that in EXO, the participants could not see the direction of *turn* directly and changed the right hand after the movement began. Furthermore, after starting to *turn* or *fold* the box, the participants changed the hand's position, presumably to ease the movement. The subjective accuracy is highest in EGO, followed by EGO & EXO and EXO, compare figure 5.4 (bottom).

carry and *walk*

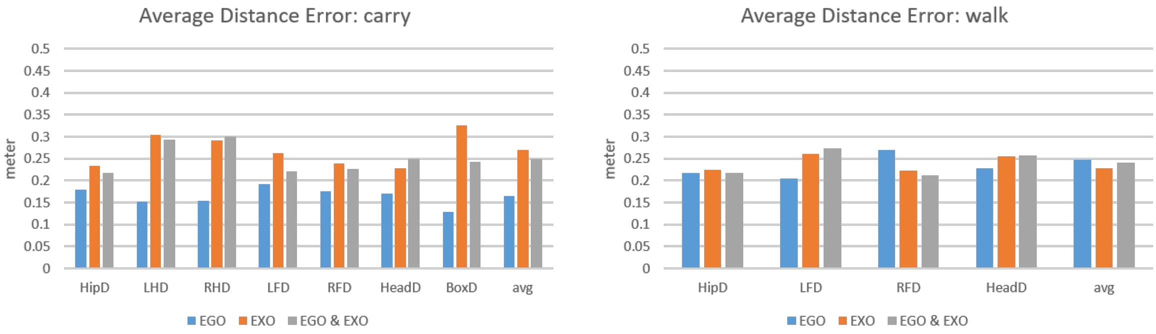


Figure 5.9.: Average error of subtask *carry* (left) and *walk* (right). Suffix D: distance. LH - left hand, RH - right hand, LF - left foot, RF - right foot.

Teaching locomotion in the ego-centric VP with the help of the *speed mechanic* is a novelty. The data revealed a nearly equal error for ego-centric guided walking and exo-centric guided walking. The position of the hands are not essential for walking and are not depicted. Adding a physical load to walking (*carry*) has a strong influence on accuracy. The learner seems to focus on the box and tries to match the GV's box with the own box. This increases the accuracy of their own position for all body parts. The subjective accuracy for *carry* is equal for EGO and EGO & EXO and higher for EXO which complies with the objective accuracy, compare figure 5.4 (bottom). The subjective accuracy for *walk* is highest in EGO.

5.5.2. Visual Focus

In EGO, the learner is provided one ego-centric GV and will focus on it. If exo-centric GVs are added to the scene, the learner can focus on multiple GVs. Furthermore, it is interesting which percentage of time the learner focuses on the own/GVs box and own/GVs body.

A pilot study helps to evaluate the experiment design and data acquisition. Conducting a pilot study before the actual study is vital. The proof is depicted in figure 5.10. In section 4.6.3 it is described how the *looking at* data acquisition method was developed and sucessfully tested. The formative test was conducted with one person, which was too little. The study data revealed that the data acquisition for the measure *looking at* is not working correctly. Over 50% of the time, the ray traces hit the environment. For the actual study, the artefacts and avatars' colliders should be adjusted, or if available, an eye-tracker should be used. Nevertheless, assuming the rays which hit something other than the environment are evenly distributed, some deductions can be made from the acquired data.

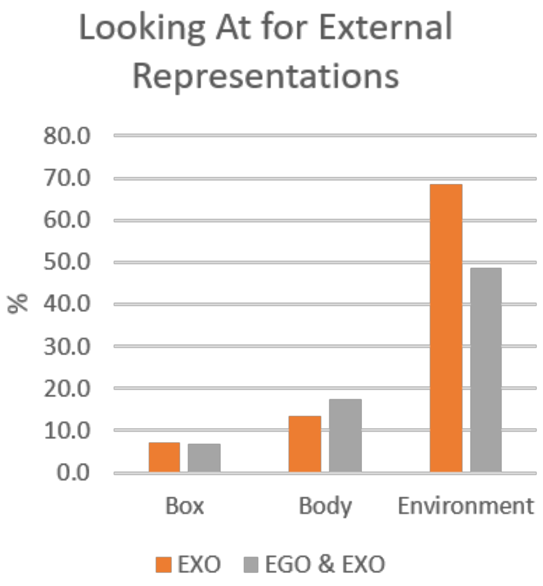


Figure 5.10.: *Looking at* for exo-centric GVs. Percentage of time focussed on the box or the avatar of the GV.

Figure 5.10 shows the percentage of time the learner focussed on a box, a body or the environment. The learner focussed roughly twice as much on a body than on the box.

Figure 5.11 shows the positions of the GVs whereby a GV is the union of body, table and box. The po-

sitions are overlayed with a heat map. The orange circles stand for the percentage of time the learner focuses on that position. The orange circles stand for the exo-centric VP, the grey circle stand for the ego- & exo-centric VP. The ego-centric VP is not depicted because there are no exo-centric GVs to share the focus of the learner. The heat map provides two insights. First, the presence of an ego-centric GV influences the visual focus of the learner. In EXO, where no ego-centric GV is present, the learner focussed 11% of the time on the artefacts (table, box) of one's own position. In EGO & EXO, the learner focussed 45% of the time on the ego-centric GV (avatar, box, table). If an ego-centric GV is present, it is more frequently focussed than exo-centric GVs. By implication, the learner is consulting the exo-centric GVs, if an ego-centric GV is present.
The second insight regards position four. In EXO and EGO & EXO, the learner did not focus on the GV at position four. Position four is superfluous for all three tasks.

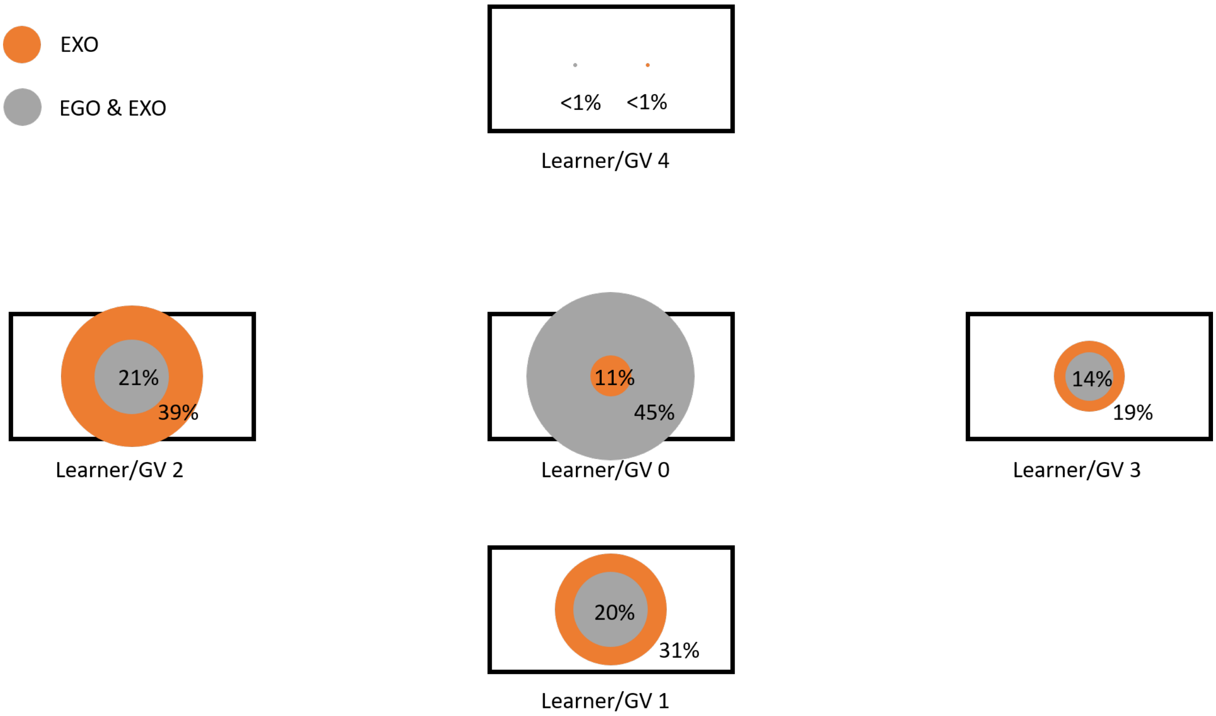


Figure 5.11.: Heat map of the learner's visual focus in EXO and EGO & EXO. The heatmap shows the tables of the GVs, compare figure 4.6. The circles' size corresponds to the amount of time the learner focussed the representation. Ego-centric VP is not depicted because in the ego-centric VP no exo-centric GV exists.

5.5.3. Risk Metrics

The risk metrics are not analysed. The reason is the missing windows the RMs are based on. Recap: for specific subtasks, the specific RM should be between a minimum and a maximum. The time inside and outside the window between the minimum and maximum yield a score. To determine a window

for the RM a professional should be consulted. Because of the COVID-19 pandemic, a determination by a professional was not possible.
However, the accuracy data for *lift* and *lower* could lead to a guess about the performance of *squat*. The feet accuracy is higher in EXO than in EGO. *Squat* refers to the same body part. Thereby, the assumption could be that *squat* could score better in perspectives with an exo-centric GV. *Support base* and *upright stance* are referring to body parts that are not directly visible in EGO, thus the assumption could be extended to *support base* and *upright stance*, too.

5.5.4. Subjective Preferences

The personal preferences of the participants tend towards perspectives with exo-centric GVs. If the participants could choose a VP for the task, two participants would decide for EGO & EXO, and one would use EXO. Additionally, two participants stated that they could follow the GV best in EGO & EXO, one could follow the GV best in EGO. All but one participant ranked the ability to follow the GV worst in EGO. This competes with the accuracy, which is the highest in EGO. Surprisingly, all participants stated that the most accessible VP was EGO. In EGO, the GV stands inside one's own body, which is not possible in real-world scenarios and thus unusual. A limitation for the ease of understanding is the high proficiency and knowledge about VR of all participants.

6. Conclusion & Outlook

This master's thesis proposes an experiment that can generate data to answer the research question RQ1: How does the visual perspective on a virtual guidance visualisation influence motor learning in Virtual Reality environments?
This chapter concludes the achieved work, improvements in experiment design and experiment implementation, and finally gives an outlook on future work.

6.1. Conclusion

In chapter 3 an experiment was described to answer the main research question. The decision to compare the ego-centric VP, augmented exo-centric VP and the ego- & augmented exocentric VP was discussed. To realise the VP, two mechanics were utilised: the *speed mechanic*, which allows ego-centric locomotion guidance and *multiple representations* which allows the learner to always see an exo-centric GV. A task that includes the handling of physical load was iteratively designed, and an experiment structure was provided. To assess the learner's performance, the experiment requires measures: accuracy measures based on distances and angles, risk metrics, focus measurements and qualitative measurements for the learner's subjective opinion. Next, chapter 4 describes the implementation of a system to conduct the experiment. The learner and teacher received a virtual representation with the help of trackers and inverse kinematics. Artefacts like the physical load were constructed and represented digitally, too. Measures required by the experiment design were implemented, and finally, the procedure of the experiment could be provided.

The experiment was evaluated with a pilot study, and vast parts of the design worked as intended. The mirror as calibration help served its purpose. Table, scale and box proved to be appropriate. The VPs were understood well by the participants. The experiment's task is safe for the participants, and the task design served its purpose for the experiment. The itinerary proved to be suited, but the total time for one participant was shortened by 10 minutes. Ego-centric locomotion guidance with the help of the *speed mechanic* works. However, adjustments for the actual experiment were necessary. The fixation of the hip tracker to the participant's body needed to be improved. An additional checklist for task and condition during the experiment were introduced. The identification of the video recordings was reconsidered. The HMD should be wireless for the actual experiment. To improve the recognition of ownership of a stationary box, GVs were suggested to be rendered with light transparency. The pilot experiment could successfully eradicate the minor errors in the log file. The method to assess the learner's focus was redesigned. Finally, three new questions were placed on the *questionnaire*. With these adjustments, the experiment is capable to generate data to answer the research question.

The pilot experiment had three participants, which does not allow a full counterbalancing. The evaluation of the data is informal. An answer to the research questions can not be given on that basis. Nevertheless, the data gives a first glimpse on what to expect from the empirical contribution of the experiment of this master's thesis. The following statements for the research questions are regarding to clear differences in the data but should still not be taken as answers. The statements are more a list of conspicuous elements in the data. The evaluation of the actual experiment can give more detailed insights.

RQ1.1 How does the visual perspective on a virtual guidance visualisation influence movements' accuracy?

RQ1.1.1 How does the visual perspective on a virtual guidance visualisation influence movements' accuracy of one's body?
The presence of an ego-centric GV seems to influence the accuracy of body parts positively. The perception of correct feet placement is limited in the ego-centric VP.

RQ1.1.2 How does the visual perspective on a virtual guidance visualisation influence the accuracy of handling physical load?
The presence of an ego-centric GV seems to influence the accuracy of the box positively.

RQ1.1.3 How does the visual perspective on a virtual guidance visualisation influence the subtasks' accuracy?
Exo-centric GVs influence the feet placement during lift and lower positively. The overall accuracy of each subtask is highest in the ego-centric VP.

RQ1.2 Does the visual perspective on a virtual guidance visualisation influence the transfer of ergonomic principles?
RM could not be evaluated. My personal subjective opinion is that the transfer of ergonomic principles could be better in VPs with an exo-centric GV.

RQ1.3 How does the visual perspective on a virtual guidance visualisation influence the learner's visual focus?
If an ego-centric GV is present, the learner focusses more on the ego-centric GV than an exo-centric GV.

RQ1.4 What is the subjective personal preference of the learner for the visual perspectives?
The participants of the pilot experiment prefer the ego- & exo-centric VP.

Based on RQ1.1-4, an assumption can be made about the main research question:
RQ1: How does the visual perspective on a virtual guidance visualisation influence motor learning in Virtual Reality environments?
The presence of an ego-centric GV increases the overall accuracy of movements for all subtasks and shifts the visual focus of the learner towards the ego-centric GV.

Yu et al. [29] published in December 2020 the guideline: use the ego-centric visual perspective if the type of motion allows, consider alternatives for other types of motions (ibid.). Based on the data of the pilot study, this guideline could also hold for full-body movements that include the handling of physical load and locomotion movements.

6.2. Outlook

E(x|g)o could benefit from being extended with the dynamic-time-warp[1] algorithm. Till now, the measures are implemented to assess the error between the learner and GV in this exact moment. The dynamic-time-warp algorithm searches in a timely window for the lowest error. This algorithm erases the reaction time. This algorithm is applied in some related work, e.g. [21], too.

Besides that, E(x|g)o is a system capable of conducting further experiments. Already, all five possible VPs, including the two which are not utilised in the proposed experiment, are already implemented. Upcoming experiments could investigate the differences between the two VP in the x-class, as well as in the gx-class.
Furthermore, E(x|g)o can be used without a physical load. Conducting a similar experiment to the proposed experiment, but without the physical load, could reveal if the outcome of both experiments correspond or differ. E(x|g)o can also be easily extended with new physical artefacts. Investigating the influence of shape, size and weight of the physical load on the learner's performance could be interesting, too. Additionally, E(x|g)o can be used for tasks where the learner is sitting. Scenarios, where the learner operates a machine seated, are possible to be evaluated with E(x|g)o.
The subtasks in the proposed experiment are all having a specific magnitude in their movement. In upcoming experiments, the magnitude can be varied.
Dürr et al. [28] showed that for ego-centric guidance, a high realistic GV tend to achieve a higher movement accuracy than abstract avatars. E(x|g)o can be used to investigate if this can also be applied to full-body avatars. Furthermore, an in-detail evaluation of the number and positions of exo-centric GV could be interesting. The distance between the learner and exo-centric GVs plays a role in how the learner can visually percept the exo-centric GVs. The influence of the distance between exo-centric GV and learner on the learner's performance is a topic of interest.
The artefact contribution of guiding ego-centric locomotion movements is limited by the low amount of participants. A larger study would erase the limitation, and the definition of the ETD_{min} and ETD_{max}

---

[1] https://towardsdatascience.com/dynamic-time-warping-3933f25fcdd, accessed 28.3.2021

can be built on empirical data. Such a study is possible with E(x|g)o too.

The generated data of the experiment proposed in this master’s thesis will help designers of VR motor learning systems to choose a reasonable perspective for their project based on empirical evidence.

References

[1] R.A. Schmidt and T.D. Lee. *Motor Control and Learning: A Behavioral Emphasis.* Human Kinetics, 2005. ISBN: 9780736042581.

[2] Fraser Anderson et al. "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. St. Andrews, Scotland, United Kingdom: Association for Computing Machinery, 2013, 311–320. ISBN: 9781450322683. DOI: `10.1145/2501988.2502045`. URL: `https://doi.org/10.1145/2501988.2502045`.

[3] J. C. P. Chan et al. "A Virtual Reality Dance Training System Using Motion Capture Technology". In: *IEEE Transactions on Learning Technologies* 4.2 (2011), pp. 187–195. DOI: `10.1109/TLT.2010.27`.

[4] Shuo Yan et al. "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. Seoul, Republic of Korea: Association for Computing Machinery, 2015, 965–970. ISBN: 9781450331463. DOI: `10.1145/2702613.2732759`. URL: `https://doi.org/10.1145/2702613.2732759`.

[5] Jacky Chan et al. "Immersive Performance Training Tools Using Motion Capture Technology". In: *Proceedings of the First International Conference on Immersive Telecommunications.* ImmersCom '07. Bussolengo, Verona, Italy: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2007. ISBN: 9789639799066.

[6] 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).* 2004, pp. 217–222. DOI: `10.1109/ROMAN.2004.1374759`.

[7] 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. Edinburgh, Scotland: Association for Computing Machinery, 2014, 55–64. ISBN: 9781450332538. DOI: `10.1145/2671015.2671023`. URL: `https://doi.org/10.1145/2671015.2671023`.

[8] Taihei Kojima et al. "Training Archived Physical Skill through Immersive Virtual Environment". In: *Human Interface and the Management of Information. Information and Knowledge in Applications and Services.* Ed. by Sakae Yamamoto. Cham: Springer International Publishing, 2014, pp. 51–58. ISBN: 978-3-319-07863-2.

[9] 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.* New York, NY, USA: Association for Computing Machinery, 2013, 1309–1318. ISBN: 9781450318990. URL: `https://doi.org/10.1145/2470654.2466171`.

[10] Richard Tang et al. "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.* New York, NY, USA: Association for Computing Machinery, 2015, 4123–4132. ISBN: 9781450331456. URL: `https://doi.org/10.1145/2702123.2702401`.

[11] Vijay Rajanna et al. "KinoHaptics: An Automated, Wearable, Haptic Assisted, Physio-therapeutic System for Post-surgery Rehabilitation and Self-care". In: *Journal of Medical Systems* 40 (Dec. 2015). DOI: `10.1007/s10916-015-0391-3`.

[12] Maurício Sousa et al. "SleeveAR: Augmented Reality for Rehabilitation Using Realtime Feedback". In: *Proceedings of the 21st International Conference on Intelligent User Interfaces.* IUI '16. Sonoma, California, USA: Association for Computing Machinery, 2016, 175–185. ISBN: 9781450341370. DOI: `10.1145/2856767.2856773`. URL: `https://doi.org/10.1145/2856767.2856773`.

[13] Maureen Holden et al. "Virtual Environment Training Improves Motor Performance in Two Patients with Stroke: Case Report". In: *Journal of Neurologic Physical Therapy* 23 (1999). URL: `https://journals.lww.com/jnpt/Fulltext/1999/23020/Virtual_Environment_Training_Improves_Motor.13.aspx`.

[14] Ping-Hsuan Han et al. "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. Geneva, Switzerland: Association for Computing Machinery, 2016. ISBN: 9781450336802. DOI: `10.1145/2875194.2875237`. URL: `https://doi.org/10.1145/2875194.2875237`.

[15] U. Yang and G. J. Kim. "Implementation and Evaluation of "Just Follow Me": An Immersive, VR-Based, Motion-Training System". In: *Presence* 11.3 (2002), pp. 304–323. DOI: `10.1162/105474602317473240`.

[16] Nicholas Katzakis et al. "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. Brighton, United Kingdom: Association for Computing Machinery, 2017, 58–67. ISBN: 9781450354868. DOI: `10.1145/3131277.3132181`. URL: `https://doi.org/10.1145/3131277.3132181`.

[17] Taku Komura et al. "E-Learning Martial Arts". In: *Proceedings of the 5th International Conference on Advances in Web Based Learning.* ICWL'06. Penang, Malaysia: Springer-Verlag, 2006, 239–248. ISBN: 3540490272. DOI: `10.1007/11925293_22`. URL: `https://doi.org/10.1007/11925293_22`.

[18] Ping-Hsuan Han et al. "My Tai-Chi Coaches: An Augmented-Learning Tool for Practicing Tai-Chi Chuan". In: *Proceedings of the 8th Augmented Human International Conference.* AH '17. Silicon Valley, California, USA: Association for Computing Machinery, 2017. ISBN: 9781450348355. DOI: `10.1145/3041164.3041194`. URL: `https://doi.org/10.1145/3041164.3041194`.

[19] Otniel Portillo-Rodriguez et al. "Real-Time Gesture Recognition, Evaluation and Feed-Forward Correction of a Multimodal Tai-Chi Platform". In: *Haptic and Audio Interaction Design.* Ed. by Antti Pirhonen and Stephen Brewster. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 30–39. ISBN: 978-3-540-87883-4.

[20] Thuong N. Hoang et al. “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. Gothenburg, Sweden: Association for Computing Machinery, 2016. ISBN: 9781450347631. DOI: `10.1145/2971485.2971521`. URL: `https://doi.org/10.1145/2971485.2971521`.

[21] Philo Tan Chua et al. “Training for physical tasks in virtual environments: Tai Chi”. In: *IEEE Virtual Reality, 2003. Proceedings.* 2003, pp. 87–94. DOI: `10.1109/VR.2003.1191125`.

[22] J. Lieberman and C. Breazeal. “TIKL: Development of a Wearable Vibrotactile Feedback Suit for Improved Human Motor Learning”. In: *IEEE Transactions on Robotics* 23.5 (2007), pp. 919–926. DOI: `10.1109/TRO.2007.907481`.

[23] 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. Austin, Texas, USA: Association for Computing Machinery, 2012, 179–188. ISBN: 9781450310154. DOI: `10.1145/2207676.2207702`. URL: `https://doi.org/10.1145/2207676.2207702`.

[24] Maximilian Dürr et al. “NurseCare: Design and ’In-The-Wild’ Evaluation of a Mobile System to Promote the Ergonomic Transfer of Patients”. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.* CHI ’20. Honolulu, HI, USA: Association for Computing Machinery, 2020, 1–13. ISBN: 9781450367080. DOI: `10.1145/3313831.3376851`. URL: `https://doi.org/10.1145/3313831.3376851`.

[25] Maximilian Dürr et al. “KiTT - The Kinaesthetics Transfer Teacher : Design and Evaluation of a Tablet-based System to Promote the Learning of Ergonomic Patient Transfers”. In: *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2021).* New York: ACM, 2021. ISBN: 978-1-4503-8096-6. DOI: `10.1145/3411764.3445496`.

[26] Stefan Paul Feyer (University of Konstanz). *Investigating Visual Perspectives on Guidance Visualisations for Motor Learning.* Seminar Thesis. 2019.

[27] Paul Milgram and Fumio Kishino. *A TAXONOMY OF MIXED REALITY VISUAL DISPLAYS.* 1994.

[28] Maximilian Dürr et al. “EGuide: Investigating Different Visual Appearances and Guidance Techniques for Egocentric Guidance Visualizations”. In: *Proceedings of the Fourteenth International Conference on Tangible, Embedded, and Embodied Interaction.* TEI ’20. Sydney NSW, Australia: Association for Computing Machinery, 2020, 311–322. ISBN: 9781450361071. DOI: `10.1145/3374920.3374945`. URL: `https://doi.org/10.1145/3374920.3374945`.

[29] X. Yu et al. “Perspective Matters: Design Implications for Motion Guidance in Mixed Reality”. In: *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR).* 2020, pp. 577–587. DOI: `10.1109/ISMAR50242.2020.00085`.

[30] Wenbi Wang. “Dynamic Viewpoint Tethering: Controlling a Virtual Camera for Effective Navigation in Virtual Environments”. In: *CHI ’01 Extended Abstracts on Human Factors in Computing Systems.* CHI EA ’01. Seattle, Washington: Association for Computing Machinery, 2001, 93–94. ISBN: 1581133405. DOI: `10.1145/634067.634124`. URL: `https://doi.org/10.1145/634067.634124`.

[31] R. Rajesh. “Manual Material Handling: A Classification Scheme”. In: *Procedia Technology* 24 (2016). International Conference on Emerging Trends in Engineering, Science and Technology (ICETEST - 2015), pp. 568–575. ISSN: 2212-0173. DOI: `https://doi.org/10.1016/j.protcy.2016.05.114`. URL: `https://www.sciencedirect.com/science/article/pii/S2212017316302031`.

[32] Jonathan Muckell, Yuchi Young, and Mitch Leventhal. “A Wearable Motion Tracking System to Reduce Direct Care Worker Injuries: An Exploratory Study”. In: *Proceedings of the 2017 International Conference on Digital Health*. DH ’17. London, United Kingdom: Association for Computing Machinery, 2017, 202–206. ISBN: 9781450352499. DOI: `10.1145/3079452.3079493`. URL: `https://doi.org/10.1145/3079452.3079493`.

[33] Winyu Chinthammit et al. “Ghostman: Augmented Reality Application for Telerehabilitation and Remote Instruction of a Novel Motor Skill”. In: *BioMed research international* 2014 (Apr. 2014), p. 646347. DOI: `10.1155/2014/646347`.

[34] Giuseppe Scavo, Fridolin Wild, and Peter Scott. “The GhostHands UX: telementoring with hands-on augmented reality instruction”. In: *Workshop Proceedings of the 11th International Conference on Intelligent Environments, Prague, Czech Republic, July 15-17, 2015*. Ed. by Davy Preuveneers. Vol. 19. Ambient Intelligence and Smart Environments. IOS Press, 2015, pp. 236–243. DOI: `10.3233/978-1-61499-530-2-236`. URL: `https://doi.org/10.3233/978-1-61499-530-2-236`.

[35] Kasper Hornbæk. “Some Whys and Hows of Experiments in Human–Computer Interaction”. In: *Found. Trends Hum.-Comput. Interact.* 5.4 (June 2013), 299–373. ISSN: 1551-3955. DOI: `10.1561/1100000043`. URL: `https://doi.org/10.1561/1100000043`.

[36] Stefan Paul Feyer. *E(x|g)o: Comparing Visual Perspectives on Guidance Visualisations for Motor Learning in Virtual Reality*. Tech. rep. Konstanz: Department of Computer Science, University of Konstanz, 2020.

[37] Maximilian Dürr et al. “EGuide : Investigating different Visual Appearances and Guidance Techniques for Egocentric Guidance Visualizations”. In: *TEI ’20 : Proceedings of the Fourteenth International Conference on Tangible, Embedded, and Embodied Interaction*. New York, NY: ACM, 2020, pp. 311–322. ISBN: 978-1-4503-6107-1. DOI: `10.1145/3374920.3374945`.

[38] Misha Sra, Aske Mottelson, and Pattie Maes. “Your Place and Mine: Designing a Shared VR Experience for Remotely Located Users”. In: *Proceedings of the 2018 Designing Interactive Systems Conference*. DIS ’18. Hong Kong, China: Association for Computing Machinery, 2018, 85–97. ISBN: 9781450351980. DOI: `10.1145/3196709.3196788`. URL: `https://doi.org/10.1145/3196709.3196788`.

A. Glossarium

VP — visual perspective

GV — guidance visualisation

MR — Mixed Reality

VR — Virtual Reality

L — learner

PT — participant

T — task

ED — Euclidean distance

ETD — ego-centric tethering distance

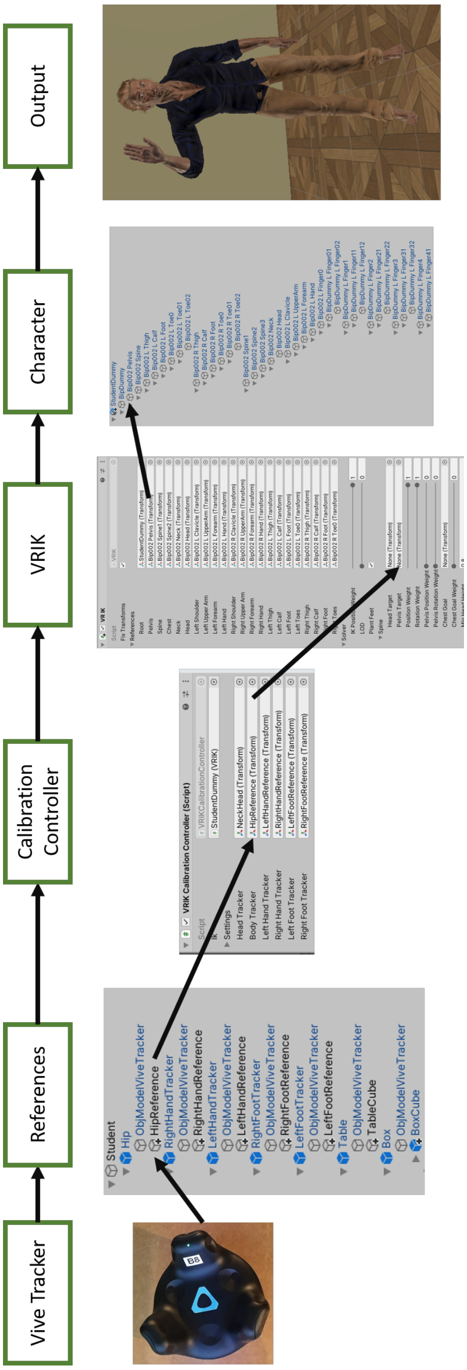
TCT — task completion time

TND — task norm duration

OTND — over task norm duration

B. Appendix

B.1. Learner Rendering Pipeline



VRIK Calibration Controller (Script)
Script VRIKCalibrationController
Ik StudentDummy (VRIK)
Settings
Head Tracker NeckHead (Transform)
Body Tracker HipReference (Transform)
Left Hand Tracker LeftHandReference (Transform)
Right Hand Tracker RightHandReference (Transform)
Left Foot Tracker LeftFootReference (Transform)
Right Foot Tracker RightFootReference (Transform)

B.2. Study Documents

B.2.1. Welcome Letter

Welcome!

Dear participant,

Thank you for agreeing to participate in this study, you are about to make a valuable contribution to my thesis! In the following you will be informed about the aim and the procedure.

You will be using a virtual reality (VR) system. As part of my master's thesis at the University of Konstanz, I created a VR application that can be used to learn movements.

To use a VR application, VR glasses are required. You will experience the virtual world in 3D. If you move with the device in the real world, you move the same distance in the virtual one. The point of a typical VR experience is to make a virtual world seem "real".

This application should now be tested by you. In VR, you will see one or more virtual teachers demonstrating a movement to you. You are supposed to copy this movement. You will interact with a 6 kg box. Your task is to follow the instructions as closely as possible. Make sure that the position of the box, your hands and legs correspond as closely as possible those of the virtual teacher. The bending of your back is also displayed ergonomically correct and should correspond as closely as possible to that of the teacher.

If the movement is lagging, check that your feet are in exactly the same place as those of the teacher. The study consists of three parts with similar tasks. However, the perspective with which you see the teacher(s) changes.

I would like to emphasize that it is the application that is being tested and not you. Follow the application as best you can, if you are too inaccurate the teacher will be waiting for you.

The procedure of the study is as follows:

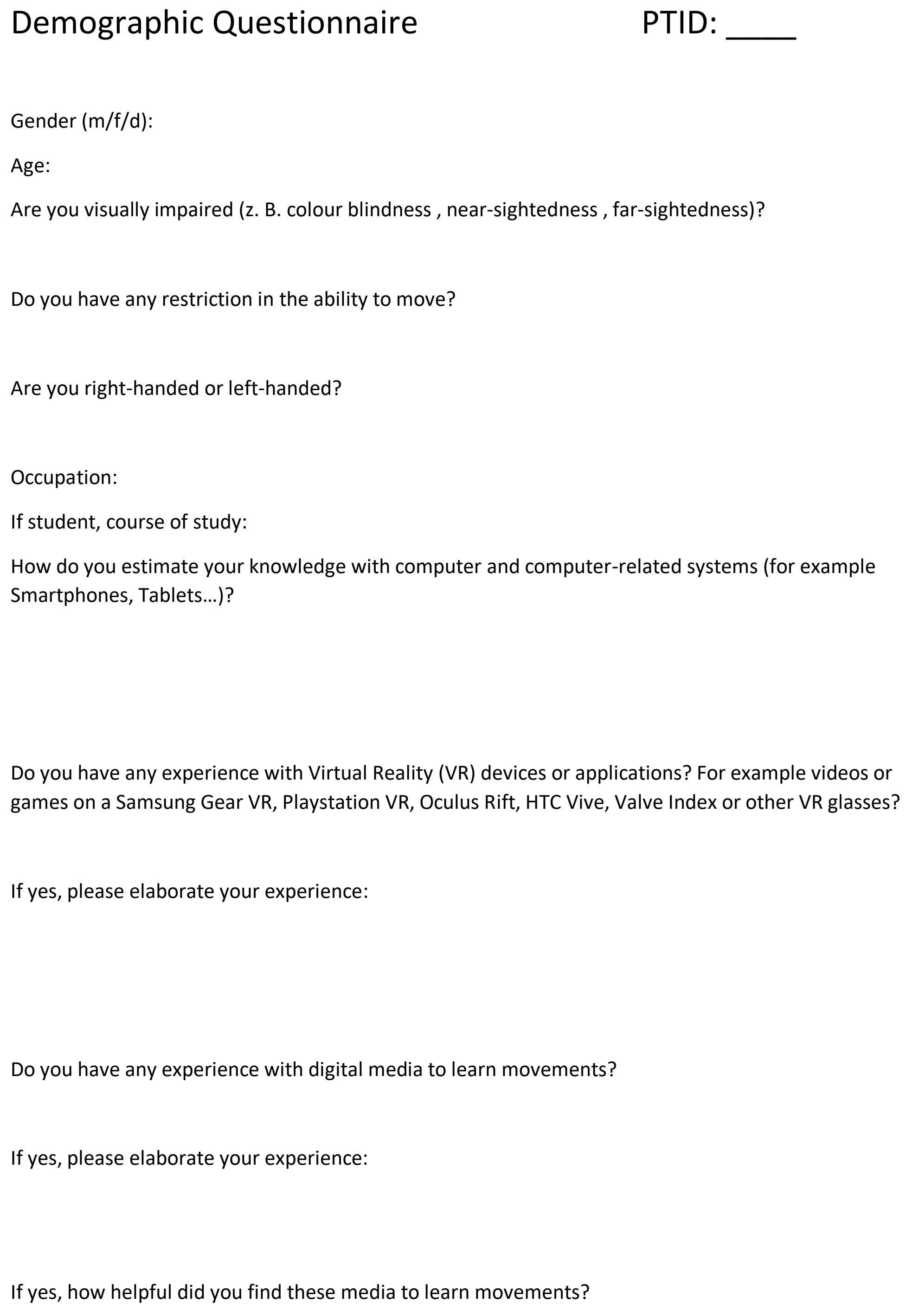
- Welcome letter, consent form, demographic questionnaire.
- Application test
- Questionnaire
- Space for comments
- Payment and receipt

The study takes about 65 minutes.

You have the option to abort this study at any time. To do so, please simply inform the study management. If you have any questions or comments, please feel free to bring them up at any time during the study!

Thank you very much for your support!

B.2.2. Demographic Questionaire



Demographic Questionnaire PTID: ____

Gender (m/f/d):

Age:

Are you visually impaired (z. B. colour blindness , near-sightedness , far-sightedness)?

Do you have any restriction in the ability to move?

Are you right-handed or left-handed?

Occupation:

If student, course of study:

How do you estimate your knowledge with computer and computer-related systems (for example Smartphones, Tablets…)?

Do you have any experience with Virtual Reality (VR) devices or applications? For example videos or games on a Samsung Gear VR, Playstation VR, Oculus Rift, HTC Vive, Valve Index or other VR glasses?

If yes, please elaborate your experience:

Do you have any experience with digital media to learn movements?

If yes, please elaborate your experience:

If yes, how helpful did you find these media to learn movements?

PTID: ____

Did you had a job where you had to move heavy loads? (For example storekeeper, craft or comparable)

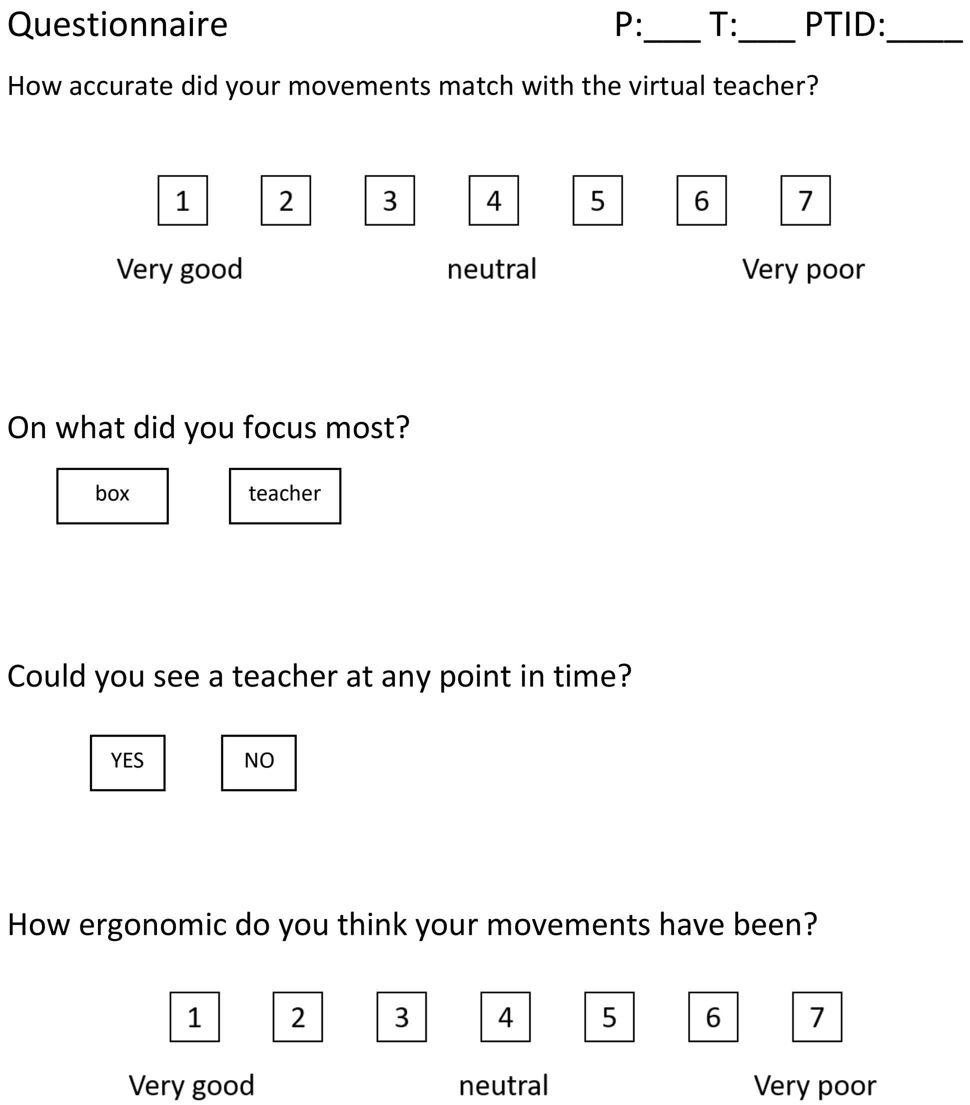
Has ergonomic conduction of handling heavy load has ever been a topic for you?

What is your knowledge about moving yourself ergonomic?

Did you ever suffer from handling heavy loads not ergonomically? (For example back pain, knee pain and comparable)

Do you try to move yourself ergonomic if you have to handle heavy loads?

B.2.3. Questionnaire



Questionnaire

P:___ T:___ PTID:____

How accurate did your movements match with the virtual teacher?

1	2	3	4	5	6	7
Very good			neutral			Very poor

On what did you focus most?

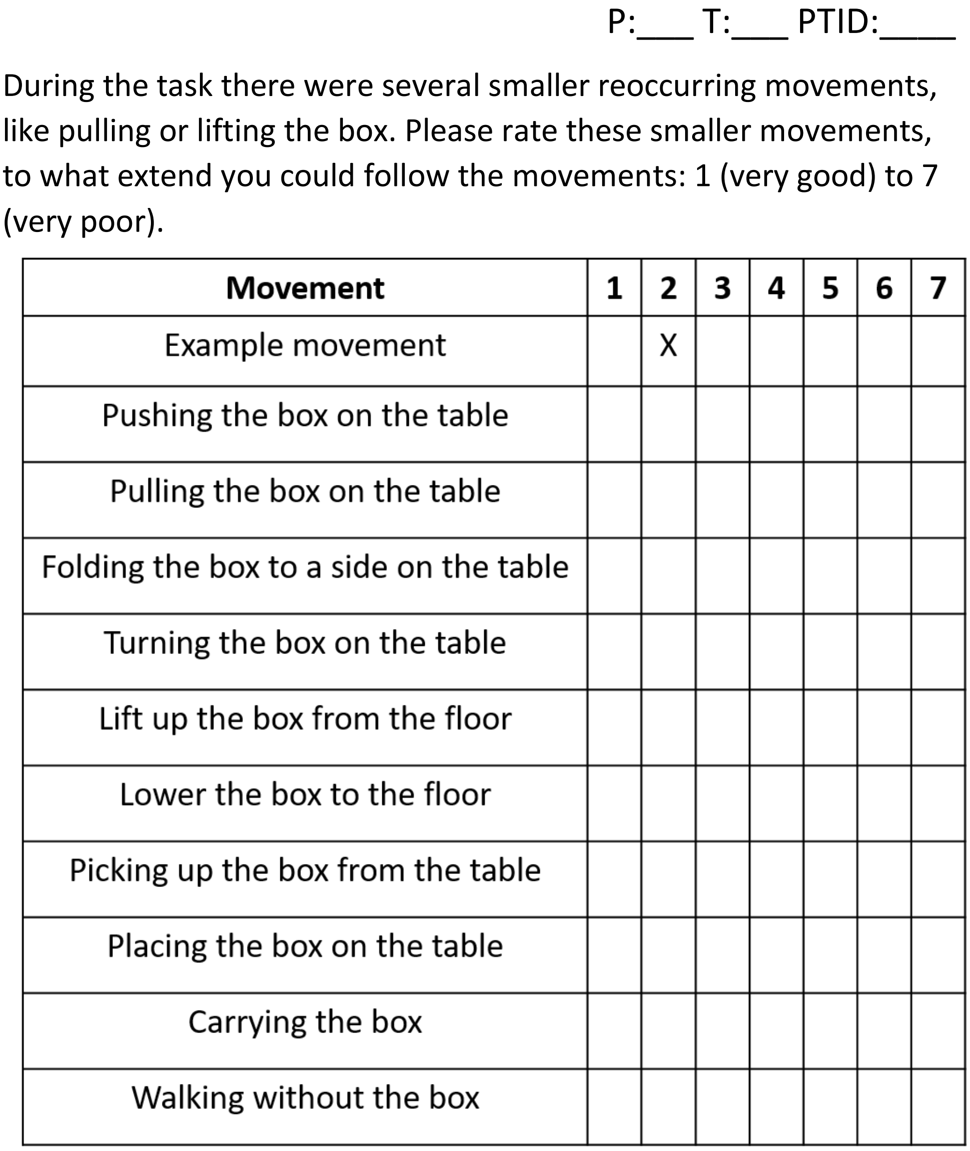
box teacher

Could you see a teacher at any point in time?

YES NO

How ergonomic do you think your movements have been?

1	2	3	4	5	6	7
Very good			neutral			Very poor



P:___ T:___ PTID:____

During the task there were several smaller reoccurring movements, like pulling or lifting the box. Please rate these smaller movements, to what extend you could follow the movements: 1 (very good) to 7 (very poor).

Movement	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Example movement		X					
Pushing the box on the table							
Pulling the box on the table							
Folding the box to a side on the table							
Turning the box on the table							
Lift up the box from the floor							
Lower the box to the floor							
Picking up the box from the table							
Placing the box on the table							
Carrying the box							
Walking without the box							

P:___ T:___ PTID:____

Please rate to what extend you think you could align your body parts with the teachers body parts: 1 (very good) to 7 (very poor).

Body part	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Example body part		X					
Legs							
Arms							
Back							

Please rate to what extend you think you could align your box with the teacher's box: 1 (very good) to 7 (very poor).

1	2	3	4	5	6	7
Very good			neutral			Very poor

P:___ T:___ PTID:____

(As interview question) Did you have problems to follow the instructions?

- E.g. because you could not see some body parts?
- E.g. bad perception related to the perspective?
- Go into extreme values of this questionnaire!
- Address critical incidences!
- Could you see a GV at any point in time?

B.2.4. Semi-Structured Interview

Semi-structured interview - guideline PTID:_____

You saw three visual perspectives: ego-centric, exo-centric and the combination. What do you think about these perspectives?

- entry question, encourage participant to talk frank, address interesting statements

Prioritise the perspectives by how accurate you could follow the movements. (1 best to 3 worst)

- Why did you prioritize this way?

Imagine you want to learn a movement in VR. Which perspective would you use for that?

- Or would you use a totally different one?

Which of the three perspectives was the easiest to understand?

- Was there a perspective you did not understand right away?
- Was there a perspective that confused you?
    - What do you think caused the confusion?

What do you think are the advantages and disadvantages of the perspectives?

PTID:_____

Could you see some body parts better or worse in the perspectives?

- What about your legs, arms, back?
- Could you detect that during *lift* and *lower* you should squat?
- Could you detect that you should step back during push and pull?

Did you miss a feature?

- Dig for improvements for E(x|g)o or experiment design.

(Space to ask for critical incidences, if any occurred.)

Digital Attachments:

- Master’s thesis
- E(x|g)o
- Experiment documents
- Experiment data
- Seminar thesis: Investigating Visual Perspectives on Guidance Visualisations for Motor Learning
- Project Report: E(x|g)o: Comparing Visual Perspectives on Guidance Visualisations for Motor Learning in Virtual Reality