

Saarland University  
Faculty of Natural Sciences and Technology I  
Department of Computer Science

**Master Thesis**

**Evaluation of Gesture-Based 3D Navigation  
Techniques in Virtual Reality**

submitted by

**Patrick Rump**

4 September 2014

Reviewers

**Prof. Dr. Jörg Siekmann**

**Prof. Dr. Christoph Igel**

## **Eidesstattliche Erklärung**

Ich erkläre hiermit an Eides Statt, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

## **Statement in Lieu of an Oath**

I hereby confirm that I have written this thesis on my own and that I have not used any other media or materials than the ones referred to in this thesis.

## **Einverständniserklärung**

Ich bin damit einverstanden, dass meine (bestandene) Arbeit in beiden Versionen in die Bibliothek der Informatik aufgenommen und damit veröffentlicht wird.

## **Declaration of Consent**

I agree to make both versions of my thesis (with a passing grade) accessible to the public by having them added to the library of the Computer Science Department.

Saarbrücken, -----  
(Datum / Date)

-----  
(Unterschrift / Signature)

# Acknowledgment

I hereby thank my supervisor Prof. Dr. Jörg Siekmann and my reviewer Prof. Dr. Christoph Igel for giving me the opportunity to write this thesis at the Center for Learning Technology (CeLTech).

I would also like to express my gratitude to everyone who supported me throughout the course of this thesis. I am thankful for their aspiring guidance, constructive criticism and friendly advices. Particular thanks go to Yecheng Gu, Patrick Walther and Nina Eberhardt.

Moreover, I would like to thank all participants for taking part in my user study.

Very special thanks go to my family for supporting and encouraging me throughout my studies.

## **Abstract**

The goal of this thesis is to design and implement several natural user interfaces using the Microsoft Kinect sensor and investigate their effect on the perceived cognitive workload during a training task. For that purpose, a dual task user study was conducted to compare three different interaction designs by measuring objective performance data as well as conducting subjective user tests. The interaction model that produces the lowest cognitive workload is integrated into the virtual reality learning application SafeChild to extend its capabilities with locomotion behavior based on a natural human computer interface.

# List of Abbreviations

<b>10y-14y</b>	Group of participants aged between 10 and 14
<b>21y-40y</b>	Group of participants aged between 21 and 40
<b>API</b>	Application Programming Interface
<b>AR</b>	Augmented Reality
<b>CeLTech</b>	Center for Learning Technology
<b>CLT</b>	Cognitive Load Theory
<b>D2V</b>	Interaction model: Distance-To-Velocity
<b>GUI</b>	Graphical User Interface
<b>HCI</b>	Human-Computer Interface
<b>HMD</b>	Head-Mounted Display
<b>IDE</b>	Integrated Development Environment
<b>IR</b>	Infrared
<b>ITS</b>	Intelligent Tutoring System
<b>LEAN</b>	Interaction model: Leaning
<b>mocap</b>	Motion Capturing
<b>NASA TLX</b>	NASA Task Load Index
<b>NUI</b>	Natural User Interface
<b>ODT</b>	Omnidirectional Treadmill
<b>SDK</b>	Software Development Kit
<b>VR</b>	Virtual Reality

**VRlow** Group of participants with low experience in Virtual Reality

**VRhigh** Group of participants with high experience in Virtual Reality

**VRmid** Group of participants with medium experience in Virtual Reality

**WiP** Interaction model: Walking-In-Place

# List of Figures

Figure 1.1	VirtuSphere, Omni-Directional Treadmill and SafeChild . . . .	2
Figure 2.1	Dance Pad controller for the Dance Dance Revolution game and D2V interaction model. . . . .	8
Figure 2.2	Examples of omnidirectional treadmills . . . . .	10
Figure 2.3	SafeChild: User approaches a crosswalk in the VR Environment.	11
Figure 3.1	The VR setup for the user study at CeLTech in Saarbrücken. . .	15
Figure 3.2	Impressions of the training level in KinectWalker . . . . .	16
Figure 3.3	Calibration of a user in KinectWalker . . . . .	16
Figure 3.4	Level designs in KinectWalker . . . . .	17
Figure 3.5	Results: NASA TLX (all) . . . . .	22
Figure 3.6	Results: NASA TLX dimensions (all) . . . . .	22
Figure 3.7	Results: Post test rating of models (all) . . . . .	23
Figure 3.8	Results: NASA TLX (age groups) . . . . .	24
Figure 3.9	Results: NASA TLX dimensions (age groups) . . . . .	25
Figure 3.10	Results: NASA TLX (gender) . . . . .	26
Figure 3.11	Results: NASA TLX dimensions (gender) . . . . .	27
Figure 3.12	Results: NASA TLX (VR Experience) . . . . .	28
Figure 3.13	Results: NASA TLX dimensions (VR Experience) . . . . .	28
Figure 3.14	Results: NASA TLX (Kinect) . . . . .	29
Figure 3.15	Results: Completion times of the main task . . . . .	30
Figure 3.16	Results: Amount of errors that were made in the main task . . .	31
Figure 3.17	Results: Time of error that was measured in the main task . . .	32
Figure 3.18	Results: Memorizing task (correct digits in prefix) . . . . .	33
Figure 3.19	Results: Memorizing task (correct digits at position) . . . . .	33
Figure 3.20	Intended walk path of the main task . . . . .	34
Figure 3.21	Walk paths of participant 3 (10y - 14y, male, VRhigh) . . . . .	35
Figure 3.22	Walk paths of participant 14 (21y - 40y, male, VRmid) . . . . .	35
Figure 3.23	Walk paths of participant 21 (21y - 40y, female, VRlow) . . . .	35
Figure 4.1	Overview of the <i>KinectWalker</i> design within a VR application. .	41
Figure 4.2	Use case diagram of <i>KinectWalker</i> . . . . .	42

Figure 4.3	The architecture of the Microsoft Kinect sensor . . . . .	44
Figure 4.4	The 20 body joints which can be tracked by the Kinect sensor. .	46
Figure 4.5	LogDog user interface within the Unity3D IDE . . . . .	49
Figure 4.6	UML diagram of the LogDog database model . . . . .	51
Figure 4.7	Sequence diagram of the update process of the virtual avatar. . .	51
Figure 4.8	Gesture definition: Rotation . . . . .	52
Figure 4.9	Gesture definition: LEAN . . . . .	53
Figure 4.10	Gesture definition: WiP . . . . .	54
Figure 4.11	Gesture definition: D2V . . . . .	54



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Problem . . . . .	1
1.2	Goal . . . . .	3
1.3	Structure of this Thesis . . . . .	3
<b>2</b>	<b>Theoretical Background</b>	<b>4</b>
2.1	Virtual Reality . . . . .	4
2.2	Cognitive Load Theory . . . . .	5
2.3	Locomotion in VR . . . . .	6
2.4	Related Works . . . . .	9
2.5	SafeChild . . . . .	10
<b>3</b>	<b>User Study</b>	<b>12</b>
3.1	Preliminary Considerations . . . . .	12
3.1.1	Hypothesis . . . . .	12
3.1.2	Design Decisions . . . . .	12
3.2	Course . . . . .	14
3.2.1	Introduction . . . . .	14
3.2.2	Practical Part . . . . .	15
3.2.3	Questionnaires . . . . .	18
3.3	Results . . . . .	21
3.3.1	Demographic Data . . . . .	21
3.3.2	Perceived Workload . . . . .	21
3.3.3	Measured Performance . . . . .	30
3.3.4	Feedback . . . . .	36
3.3.5	Conclusion . . . . .	36
<b>4</b>	<b>Application - KinectWalker</b>	<b>39</b>
4.1	Concept . . . . .	39
4.1.1	Requirements . . . . .	39
4.2	Design . . . . .	40
4.3	Implementation . . . . .	43

4.3.1	Microsoft Kinect Sensor . . . . .	43
4.3.2	Used Software . . . . .	45
4.3.3	LogDog . . . . .	48
4.3.4	Interaction Models . . . . .	50
4.4	Integration into SafeChild . . . . .	55
<b>5</b>	<b>Conclusion</b>	<b>56</b>
5.1	Summary . . . . .	56
5.2	Findings . . . . .	57
5.3	Outlook . . . . .	58
5.3.1	Evaluation . . . . .	58
5.3.2	Implementation . . . . .	58
<b>A</b>	<b>Appendix</b>	<b>59</b>
A.1	Pre Test Questionnaire . . . . .	59
A.2	Email Response Sheet . . . . .	60
A.3	NASA TLX Questionnaire . . . . .	61
A.3.1	First part of NASA TLX . . . . .	61
A.3.2	Second part of NASA TLX . . . . .	62
A.3.3	Information Sheet . . . . .	63
A.4	Post Test Questionnaire . . . . .	64
A.5	Results of the User Study . . . . .	65
A.6	Walk Paths . . . . .	72
	<b>References</b>	<b>78</b>

# 1. Introduction

Virtual Reality (VR) technology has evolved over the last decade and has become a valuable tool that is widely used in various research fields but also in the consumer market. Especially in the field of education and training, VR systems have successfully been used in the past (e.g., Pletcher, Bier, & Lubitz, 2000; Roussos et al., 1997). Their ability to simulate realistic and safe environments allows developers to create training applications in which users may encounter situations that would e.g. harm themselves or others in similar real world scenarios.

Today, low-cost consumer products like stereoscopic 3D monitors, Head-Mounted Displays (HMD) or 3D full body tracking sensors like the Microsoft Kinect sensor<sup>1</sup> allow more developers to design efficient natural user interfaces (NUI), which can increase the usability of their applications. An efficient “NUI relies on a user being able to carry out natural motions, movements or gestures and let them quickly discover how to control the computer application” using gestures or speech in a direct and intuitive manner (Bin Tomi & Rambli, 2011). The importance of such systems is also proven by the acquisition of the company Oculus VR by Facebook for \$2 billion in 2014 even before the company has released the final version of its HMD, called Oculus Rift<sup>2</sup> (Kovach, 2014).

## 1.1. Problem

Immersive VR systems have always been a reliable tool for researchers to create realistic training setups for different groups of users, addressing e.g. medical (Pletcher et al., 2000), industrial (van Wyk & de Villiers, 2009) or military (Darken, Cockayne, & Carmein, 1997) needs. With the release of the Microsoft Kinect sensor in November 2010 and similar devices afterwards, VR systems could be enriched with low-cost full body gesture-controlled interfaces. Comparing this to complex and expensive setups like VirtuSphere (Medina, Fruland, & Weghorst, 2008) or setups with omnidirectional treadmills (Darken et al., 1997) (cf. fig. 1.1), the Kinect sensor and similar devices can be seen as a disruptive technology. Since VR devices like this became affordable to nearly everyone in the last years, VR systems expanded from the research area to the

---

<sup>1</sup><http://www.microsoft.com/en-us/kinectforwindows/> (last accessed 2 August 2014)

<sup>2</sup><http://www.oculusvr.com/> (last accessed 2 August 2014)

consumer market and also to the field of education. Nevertheless, most systems are still expensive, making it hard to find training setups for low budget institutions like schools.

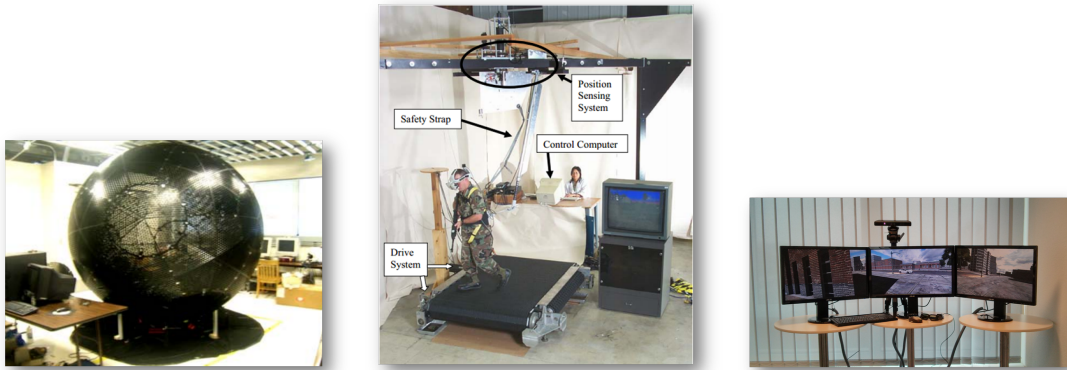


Figure 1.1.: On the left: The VirtuSphere setup (Medina et al., 2008). In the center: Omni-Directional Treadmill (Crowell III et al., 2006). On the right: SafeChild setup (Gu, 2014)

At the German Research Center for Artificial Intelligence (DFKI), an ongoing project called SafeChild (Gu & Sosnovsky, 2014) tries to build up a VR system that offers an intelligent learning system for young children in order to teach them fundamental pedestrian safety rules. It is based on a triple screen setup as output and the Microsoft Kinect sensor as input device to provide a marker-less tracking of users and to enable the users to move freely within a virtual environment. However, there is no universal definition for an effective NUI for the purpose of locomotion behavior in a training setup when using the Kinect sensor. There is also no information available that describes the influence of such NUIs on the cognitive capacities of the users. Therefore, the correlation between different NUI implementations, based on the Microsoft Kinect sensor, and their influence on a user's cognitive capacities within a VR training setup will be investigated in this thesis. Furthermore, the input capabilities of the SafeChild project should be extended by integrating one well-designed NUI based on the findings of this thesis.

## 1.2. Goal

The goal of this thesis is to provide new findings about how different gesture-based NUIs in the context of locomotion in VR applications can be used for VR training systems. For that purpose, this thesis discusses, investigates and evaluates such NUIs focusing on real life locomotion behavior.

Based on an introductory discussion about feasible NUIs, a set of promising interaction models will be selected to investigate the cognitive load (Paas, Tuovinen, Tabbers, & Van Gerven, 2003) they produce and their overall usability in correlation to their level of abstraction. This investigation will be done by conducting a user study in order to find out which model produces the lowest cognitive load while still providing a high usability for the user of a VR system. The most promising model from the user study will then be integrated into the SafeChild project (Gu & Sosnovsky, 2014) as a prototype implementation.

For the purpose of testing the NUIs, a VR application called *KinectWalker* was developed, which integrates different NUI implementations and which is able to let users test these NUIs in a pre-defined test scenario. It is further able to visualize performance data which has been recorded during the execution of a test. Each of the NUIs, which are implemented in the *KinectWalker*, are defined by one interaction model and only make use of the Microsoft Kinect sensor as physical input device.

## 1.3. Structure of this Thesis

This thesis is structured as follows: In **chapter 2**, the theoretical background for the understanding of the upcoming chapters is given. It introduces the concept of *Cognitive Load Theory* in the context of this work and explains the term *natural user interface* in the context of VR. Furthermore, a discussion about promising interaction models concerning locomotion in 3D worlds is given and related works are presented. **Chapter 3** describes the design, the execution and the evaluation of the user study of this thesis. After that, **chapter 4** describes in detail the VR application *KinectWalker*, which has been used in the user study. This chapter describes the concept and design of the application and gives detailed technical information about the main parts of its implementation. The last part of chapter 4 describes the integration of the *KinectWalker* into the SafeChild project. Finally, **chapter 5** sums up the objectives and results of this thesis and gives an outlook on future work.

For reasons of clarity and comprehensibility the words of one gender in this thesis include both other genders.

## 2. Theoretical Background

This chapter will introduce the reader to the theoretical background of this thesis. As there is no unique definition for the term *Virtual Reality*, the first section will explain how the respective term is used in the context of this work. Moreover, the main ideas of the Cognitive Load Theory of Chandler and Sweller (1991) will be presented to provide the knowledge base for the understanding of the user study from chapter 3. Finally, some related works will be presented.

### 2.1. Virtual Reality

Although not clearly defined in the literature, the term *Virtual Reality* (VR) is commonly used to refer to the variety of existing immersive systems. In 1965, Ivan Sutherland introduced the first Head-Mounted Display (HMD) that was able to show a cube's wireframe (Poupyrev et al., 2002). While his invention rather belongs to the field of *Augmented Reality* (AR), it can also be referred to one of the first human-computer interface devices in VR since AR was oftentimes used similar to VR in former times. Today, the technique of using HMD in VR systems is still present, as the example of the Oculus Rift HMD shows.

According to Rogers, Sharp, and Preece (2007), VR uses computer-generated, graphical simulations in order to create the illusion of being present in a synthetic environment. While this description mainly focuses on the graphical output of VR systems, the definition of Burdea and Coiffet (2003) provides a definition of VR that covers several aspects at once. They say that VR “is a high-end user-computer interface that involves real-time simulations and interactions through multiple sensorial channels” (Burdea & Coiffet, 2003). As this definition defines VR to be a user-computer interface that addresses multiple sensorial channels of a user, this thesis will also refer to this definition when using the term VR.

#### Natural User Interface

Human-computer interaction is a research field which “aggregates a collection of semi-autonomous fields of research and practice in human-centered informatics” (Carroll, 2009), like computer science, cognitive psychology, media studies and others. It covers

the study, the design and the usage of the interaction between humans and computers in order to improve its usability. A human-computer interface (HCI) describes the point at which a user directly communicates with a computer system. In addition, a NUI describes HCIs in which users can interact with a system by speech, touch or gestures-based interactions. The goal of an effective NUI is to enable users to interact intuitively with a system based on existing knowledge from everyday behavior. In the context of this thesis, *walking* refers to such an interaction, for which several NUIs will be evaluated.

## 2.2. Cognitive Load Theory

The Cognitive Load Theory (CLT) describes a theory about the cognitive load in the context of learning. For the purpose of this thesis, it is important to know that sensorial perceived information is processed within the so-called working memory, which has only a “very limited capacity when dealing with new information” (Artino Jr, 2008). For this reason, it is necessary to reduce the amount of non-relevant information being perceived during a learning task in order to achieve effective learning results. Thus, reducing as much of the cognitive load in the working memory as possible is desirable.

According to CLT, there are three different kinds of cognitive load:

- **Intrinsic cognitive load:** Describes the number of elements that have to be processed simultaneously in the working memory in order to create new information which can be saved in the long term memory (Artino Jr, 2008). If e.g. a large number of vocabulary must be learned, intrinsic cognitive load is reserved in order to process the amount of new words.
- **Extraneous cognitive load:** This kind of cognitive load describes the progress of processing information which is not relevant for the intended learning material. In addition, extraneous cognitive load is said to be additive to intrinsic cognitive load and, therefore, influences learning processes negatively by blocking cognitive resources in the working memory (Artino Jr, 2008). Regarding the example of learning new vocabulary, extraneous cognitive load is reserved for tasks like searching synonyms in a dictionary.
- **Germane cognitive load:** Germane cognitive load is responsible for combining perceived information and extend existing knowledge. Thus, if intrinsic and extraneous cognitive load are small enough, germane cognitive load is reserved for processes that are directly relevant for learning (Artino Jr, 2008). In the case of learning new vocabulary, germane cognitive load is reserved to create helpful

structures like linking words to familiar pictures or creating stories out of the new words that help learners memorize them.

In addition to the general concept of CLT, Jacobson (2008) said that the sensorial input channels of a person are also limited in their capacity. That means that different kinds of sensorial information can be perceived over different input channels while they must then be processed within the working memory at the same time.

This thesis will investigate the amount of extraneous cognitive load that is produced while navigating through a virtual world using different natural interaction models for locomotion.

### 2.3. Locomotion in VR

Human beings in real life have a fundamental knowledge of their own presence in the world. That is caused by a continuous and unconscious sensory flow from the moveable body parts and their relation to each other (Slater, Usoh, & Steed, 1995). This data flow is called proprioception and allows us to create a *mental model* of our body. e.g. when clapping one's hands together with closed eyes (Slater et al., 1995). In a virtual world, however, a person is often represented by a virtual avatar that should mimic real life movements. Slater et al. (1995) refer to this self-representation as the *virtual body*. They pointed out that the effectiveness of a VR system also depends on the consistency between the user's *mental model* and the *virtual body*. Thus, VR systems "offer the possibility of perceptually immersing individuals into computer-generated environments, and yet the typical means for the most basic form of interaction — locomotion — do not at all match the physical actions of walking in reality" (Slater et al., 1995).

Moreover, the implementation of a NUI, which provides locomotion interaction as it is intended in this thesis, always depends on the VR system in which the NUI is integrated. Therefore, the following aspects were considered for choosing the interaction models that are tested in the presented user study.

- **Space:** As a stationary VR setup is used for conducting the user study of this work, space was limited by the field in which the Kinect sensor is able to operate. This is crucial since real walking is endless in space and only limited by physical obstacles. Stated another way, "it is not possible to navigate an infinite VE within a finite VR system using completely natural locomotion techniques" (Marsh, Putnam, Kelly, Dark, & Oliver, 2012). Caused by a limited space, the walking speed must also be abstracted by the selected interaction model.
- **Hard- and software:** This thesis uses a hardware setup which is based on the Kinect sensor as input device to recognize full-body gestures. However, the de-



vice is limited in terms of the field in which it can operate reliably. It also depends on the resolution of the captured color and depth images of the sensor how accurate gestures can be recognized. For the purpose of this thesis, interaction models are therefore required to be based on easy and distinguishable full-body gestures.

In addition to these technical aspects, “the powerful illusion of immersion may be lost through naive interaction metaphors borrowed from non-immersive forms of human-computer interaction” (Slater et al., 1995).

Three interaction models were implemented for the user study of this thesis, focusing on the above-given constraints. In the following, they will be introduced and evaluated respecting the assumed level of abstraction compared to real walking behavior. A detailed and technical description of the implementation is given in chapter 4.

### **Lean**

The Lean model (LEAN) is based on the interaction that is known from common joystick devices. This implies that a leaning gesture will result in an according movement within the virtual world either into a forward or backward direction. Sideways movements are not supported since they are assumed to be too exhausting for the users.

**Advantage:** The model provides a simple interaction to move forward and backwards. It also simulates speed according to the angle of the corresponding leaning gesture.

**Disadvantage:** The model does not address real walking gestures since it only uses upper body parts for interaction. When users move forward fast, it is also difficult to focus the screen of the VR setup. In addition, constantly leaning forward and backwards probably exhausts the users.

Regarding the level of abstraction from real walking behavior, LEAN is assumed to be the most abstract of the selected models since it lacks spatial movements as well as gestures that are based on real walking gestures. LEAN is therefore assumed to produce the highest cognitive workload for users because it is less intuitive and needs more cognitive capacity.

### **Walking-In-Place**

The Walking-In-Place (WiP) model is defined by two simple gestures. Lifting one knee after the other while standing on one spot in the room instructs the system to move forward. Speed can be controlled by increasing or decreasing the time between lifting the left and right knee alternately.

**Advantage:** The model mimics real walking and should be easy to understand. Moreover, users can control the speed of virtual movements.

**Disadvantage:** The model provides inconsistency in terms of spatial movement since the user only performs gestures in one position without moving around in real life, while the virtual avatar in a VR application will move forward. In addition, the model might cause frustration as the sensor fails to recognize some lifting gestures. Moreover, walking backwards is not possible.

Regarding the level of abstraction from real walking behavior, WiP is assumed to be less abstract than LEAN. Even though it lacks spatial movements, it tries to address real walking gestures by a similar interaction of lifting one knee after the other on the spot. It is therefore assumed to produce less cognitive workload than LEAN in the user study.

### Distance-To-Velocity

The model Distance-To-Velocity (D2V) originates from the input device of the Dance Dance Revolution game (cf. left image in fig. 2.1). The original device consists of a rectangle pad, which is placed on the floor with additional arrow buttons. These buttons point into the four directions: forward, backwards, left and right.

The D2V model, however, extends the original dance pad control by an infinite number of possible directions based on a circular interaction field (cf. right image in fig. 2.1). Therefore, the user is able to move freely within a circular interaction field. In addition, speed can be controlled by the distance that a user moves into one direction.



Figure 2.1.: On the left: the original Dance Pad controller for the Dance Dance Revolution game (Wikipedia). On the right: the extended controller function for the D2V interaction.

**Advantage:** The user can freely walk around in the VR setup, which is desirable for immersive VR systems. Moreover, the user simply moves into the direction into which he wants the virtual avatar to move. Even backwards and sideways movements are pos-

sible. Finally, speed is simulated according to the moved distance of the user from the center of the interaction field.

**Disadvantage:** Since speed is proportional to the moved distance, stopping cannot be performed naturally because the user must first return to the center of the interaction area. The model is further limited in functionality due to the size of the available interaction area.

Concerning the level of abstraction from real walking behavior, D2V is assumed to be the most natural model from the selected ones. It tries to copy real walking behavior by letting the user walk freely around within a defined interaction area. However, it abstracts in terms of speed. Nevertheless, it is assumed that D2V will perform best in the user study and, therefore, produces the lowest amount of cognitive workload.

## 2.4. Related Works

In the following, some related works will be presented, which introduce a new way for natural locomotion interaction in VR systems, or which conduct user studies to compare different interaction techniques.

### Omnidirectional Treadmills

An omnidirectional treadmill (ODT) is a concept for VR devices that provides a surface on which a person can walk naturally without leaving the surface.

Today, several researchers and companies work on devices based on the ODT concept. The most popular one is *Virtuix Omni*, which was developed in the context of a Kickstarter campaign. It is a gaming system based on a concave platform, which enables immersive walking and running interaction (cf. fig. 2.2). For the purpose of walking in place, a low-friction surface and special shoes are used. Besides the *Virtuix Omni*, the ODT of Crowell III et al. (2006) (cf. fig. 1.1) or the *WizDish* devices of Swapp, Williams, and Steed (2010) (cf. fig. 2.2) describe similar approaches, to name only a few. In any case, such systems are still expensive and not as portable as the Kinect sensor.

### A Study about Navigation and Interaction Techniques

In their work, Dam, Braz, and Raposo (2013) conducted a study in order to compare different techniques for navigation and interaction in VR environments using the Kinect sensor. They investigated the techniques in terms of usability and performance. For the



Figure 2.2.: Examples of omnidirectional treadmills. On the left: *Virtuix Omni* (kickstarter.com). On the right: the *Wizdish* input device (wizdish.com).

selection of feasible techniques, they focused on combinations of navigation and selection interaction.

For navigation, they introduced the techniques *virtual foot*, which can be described by a virtual version of the before-mentioned dance pad, *Dial Pads*, which is a hand-based interaction control as it is used by mobile games nowadays, and *Virtual Circle*, which is a technique similar to the D2V model of this thesis.

Taking only their findings for the navigation task into account, *Virtual Circle* provided the highest interaction fidelity compared to the other two. However, parameters like the level of abstraction or the demographic data of their participants were not considered.

## 2.5. SafeChild

As one of the introduced interaction models will be integrated in the SafeChild project of Gu and Sosnovsky (2014), this project will be introduced in the following.

SafeChild is a project “which aims at the overall development of a modern virtual training system for child pedestrians.” (Gu & Sosnovsky, 2014). The project is part of the *Software Campus* and is funded by the *Bundesministerium für Bildung und Forschung* (BMBF) in cooperation with the *Center for Learning Technology* (CeLTech). It is further supported by the companies *Deutsche Post AG* and the *Landesvermessungsamt of Saarbrücken*.

Being immersed in a VR environment, young students train their theoretical knowledge in road safety education practically by performing realistic and dynamic training tasks. Performing these tasks ought to increase the student’s knowledge and skills in the domain of child pedestrian safety. One exercise is e.g. crossing a street with high traffic

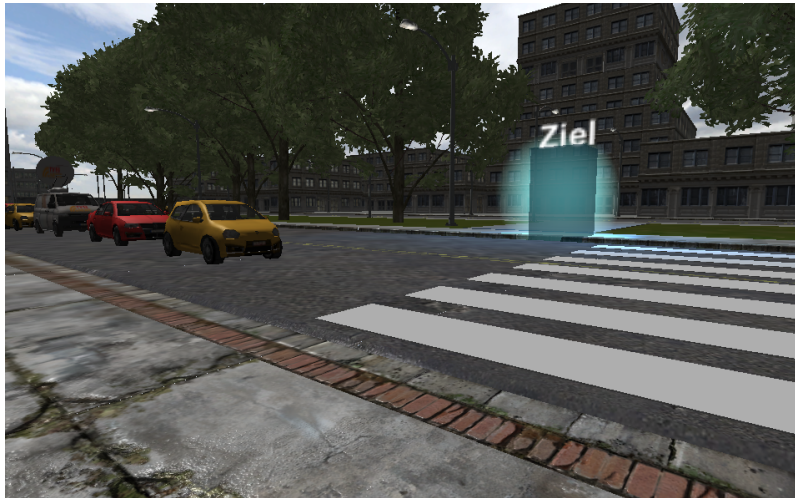


Figure 2.3.: SafeChild: User approaches a crosswalk in the VR Environment.

density by using a crosswalk (see fig 2.3). During this exercise, the student has to approach the crosswalk and check the street for approaching cars. The student must cross the street using the crosswalk when he has ensured that the situation is safe. As input device the Microsoft Kinect sensor is used to provide a natural and intuitive interaction.

One of the core components of the integrated intelligent tutoring system (ITS) in the SafeChild project is its intention recognition model, which tries to derive the user's intention in order to diagnose causes for errors or other observed behaviors. An ITS is a system that should “engage the students in sustained reasoning activity and [...] interact with the student based on a deep understanding of the student's behavior” (Corbett, Koedinger, & Anderson, 1997). Therefore, “even the basic locomotion could indicate training-related intentions” (Gu & Sosnovsky, 2014).

## 3. User Study

This chapter will describe the user study, which has been conducted in the context of this thesis. It will provide answers to the question in how far different interaction models affect the cognitive performance during a training scenario. It will also show how demographic parameters of the participants may influence the results.

### 3.1. Preliminary Considerations

#### 3.1.1. Hypothesis

As discussed in Chapter 2, the three promising interaction models *Leaning* (LEAN), *Walk-In-Place* (WiP) and *Distance-To-Velocity* (D2V) were selected as promising natural interaction possibilities for the purpose of locomotion in a 3D environment. According to the goals from section 1.2, the following hypotheses were formulated:

- H<sub>1</sub> *The level of abstraction of an interaction model compared to real walking behavior affects the perceived cognitive load of the participants. The more natural a model is, the lower the workload.*
- H<sub>2</sub> *The demographic parameters “gender” and “age” of a participant have a significant effect on the perceived workload.*
- H<sub>3</sub> *Participants without or with less experience in VR will show a lower cognitive performance than participants with VR experience.*
- H<sub>4</sub> *The model that provides the best results in the memorizing task also outperforms the others in the training scenario considering “completion time” and “time of error”.*

#### 3.1.2. Design Decisions

The user study was primarily conducted at the Center for Learning Technology (CeL-Tech) in Saarbrücken to ensure that each participant of the study uses the same setup in the same physical environment. That provided a stable and controllable setup for the course of the study. In addition, it could be ensured that the exercises were understood

and performed correctly by all of the participants and help could be provided to those participants who had problems with filling out the complex questionnaires. A similar setup to the one in Saarbrücken was build up in Trier, too, since some of the participants were not able to come to Saarbrücken.

To answer the questions that are proposed in this thesis, it was necessary to have as many young participants as possible for this study and at least the same number of teenagers or adults between the ages of 15 and 20 or 21 and 40. Except from the parameter *age*, the participants were selected randomly.

A training session was defined by a pre test questionnaire to gather demographic data from the participants, several practical tasks with a questionnaire for each of the tasks and, finally, a post test questionnaire.

### **Dual Task**

The design of the practical tasks of the user study followed the dual task paradigm from the field of psychology research (cf. Shallice, McLeod, & Lewis, 1985). This paradigm describes a procedure in which an experimentee has to perform two tasks simultaneously. He is supposed to concentrate on both tasks in the same way by trying to reach the best results possible for him. This kind of task made it possible to combine a navigation task that addresses the physical interaction with the VR setup as well as a cognitive demanding memorizing task, the results of which were used as an objective measurement for the cognitive performance of a participant. In detail, the goal of the navigation task was to reach several landmarks within a virtual world by using one of the selected interaction models. At the same time, the participants had to memorize a given number of nine digits in the right order. The amount of nine digits was selected on purpose, considering the  $7\pm 2$  rule of Miller (1956). It should be ensured that the short-term memory of the participants was filled up to its average capacity of seven digits.

The selected tasks should not interfere with each other when information is perceived and transferred to the working memory. As Jacobson (2008) has pointed out, also the input channels of a person have only a limited capacity when information is received in the context of learning. Thus, the navigation task should primarily address the visual input channels while the digits for the memorizing task were read out loud to the participants to address their aural input channel.

### **KinectWalker**

In order to test the three selected interaction models and to measure performance data during a dual task exercise, a framework called *KinectWalker* was developed. *Kinect-*

*Walker*, in general, is a program that provides a simple possibility of developing and integrating different kinds of interaction models in a VE using the Microsoft Kinect sensor. It basically adds locomotion capabilities which are based on full body tracking techniques to any pre-designed 3D world using the game engine Unity3D<sup>1</sup>. In addition, performance data like the completion time of tasks or the locations of a participant can be logged during a training session by the integrated logging application *LogDog*. Further information about the applications *KinectWalker* and *LogDog* and a more technical description of the implementation is given in chapter 4.

## 3.2. Course

In this section the detailed course of the user study will be presented. It will be shown what each of the participants had to do during a session and what kind of questionnaires they had to fill out. A session, in general, covers an introduction of the study, a practical part, in which participants test different NUIs, and three kinds of questionnaires.

### 3.2.1. Introduction

Most of the participants were acquaintances or colleagues. For those participants younger than 18 years, a parental permission was asked for in advance, which included a short description of the study.

First, the goals of the study were explained to the participants. The way of explaining the goals of the study was adjusted to the age of the participant. Young participants were told that there are people who are currently developing a computer program to teach traffic rules to children and that the participants were performing the training session to help these people finding the best input control for this program. This should make the children feel important about their participation in this study. They were also told that they could not do anything wrong and that everything they said or thought could help reach the presented goals. Adult participants got a more technical and detailed description of the goals of the study.

After the introduction, the VR setup (see fig. 3.1) was presented to the participants and the pre test questionnaire was handed out (see appendix A.1). For young participants, the questions were read out loud and the answers were then filled in by the director of studies. The participants were offered the option of writing down their E-Mail address for being informed about the results of this study later on (see appendix A.2). This option was used by four participants.

---

<sup>1</sup><http://unity3d.com/> (last accessed 1 August 2014)



### 3.2.2. Practical Part

After the introduction and the pre test questionnaire, the participants performed the practical exercises of the study. They were told to stand in the middle of the interaction field of the setup, which was an equilateral octagon around a cross on the ground (see fig. 3.1). The cross was located approximately 2,5 meters in front of the Kinect sensor and should help the participants keep themselves centered to the sensor while the octagon indicated the approximate size and form of the circular interaction area. When the participant was located at this initial position, the *KinectWalker* application was configured and started. For each of the three selected interaction models, the following two exercises had to be performed. After both exercises a questionnaire had to be filled out before continuing with the next interaction model.

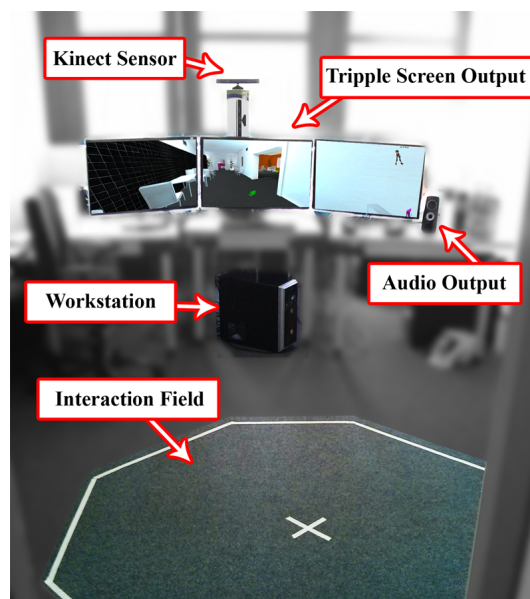


Figure 3.1.: The VR setup for the user study at CeLTech in Saarbrücken.

#### Training Exercise

The first task for the participant was to navigate through a house and collect ten landmarks. The landmarks were designed as animated circles of light (see fig. 3.2 on the left), which vanished as soon as the participant walked through them. Each landmark was shown individually so that each participant had to collect the landmarks in the same order. A green arrow at the center screen pointed into the direction of the next landmark (see fig. 3.2 on the right). The participant was also told to avoid running into obstacles like flowers, walls, tables or touching anything else than the floor (see fig. 3.2 in the center). This first exercise was not relevant for the performance which is evaluated at

the end of this chapter. It should rather introduce the participant to the selected interaction model and allow him to experience the feeling of how to navigate through the 3D world.

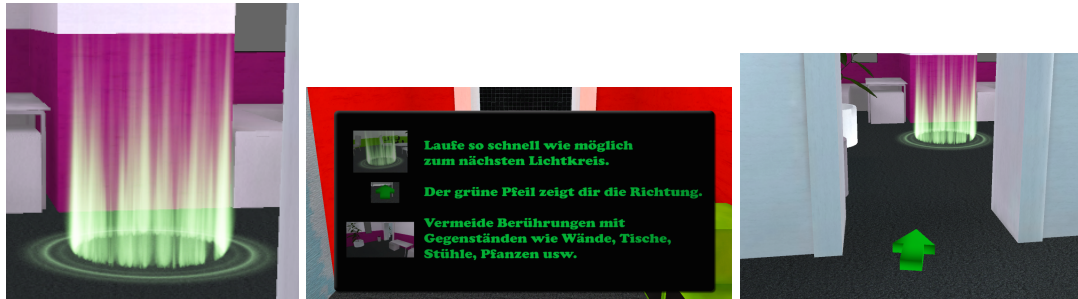


Figure 3.2.: Impressions of the training level in *KinectWalker*. From left to right: a landmark in form of a circle of light; a menu to inform the participants about their task; an arrow that shows the direction to the next landmark.

Before the participant started the training exercise, he was presented the selected interaction model. The participant was shown how to control the movements and uncertainties were resolved. As soon as the participant understood the model, he could initiate the exercise by performing the *PSI pose* (see fig. 3.3 on the left) to start the calibration procedure. After two seconds of calibration, the measured body height of the user was computed and the height of the camera in the virtual world was adapted accordingly (see fig. 3.3 on the right). This adjustment should achieve a more natural view angle for the participant than a hard-coded height value because the participant was now able to experience the virtual world from a similar perspective as he was used to in the real world. A simple waving gesture was used to start the exercise when the participant was ready.



Figure 3.3.: Calibration in *KinectWalker*. On the left: a virtual avatar of the Unity3D game engine shows how to perform the PSI calibration pose. On the right: the body height is displayed after the calibration process has been finished.

## Main Exercise

After the completion of the training exercise, the participant was assumed to feel familiar with the interaction model and the main exercise was initiated by the application. In this main exercise, the participant was given two tasks, on which he should concentrate at the same time and with the same priority.

The first task was to collect ten landmarks as fast as possible similar to the training exercise. The number of landmarks was not mentioned and the level design also differed from the one in the training exercise (see fig. 3.4). Moreover, the green arrow was not displayed in the main task in order to force participants to search for the next landmark by themselves. However, the level design was simpler and successive landmarks were located near each other. As a further constraint, the participants were told that leaving the green path would count as an error and that they should commit as few errors as possible.

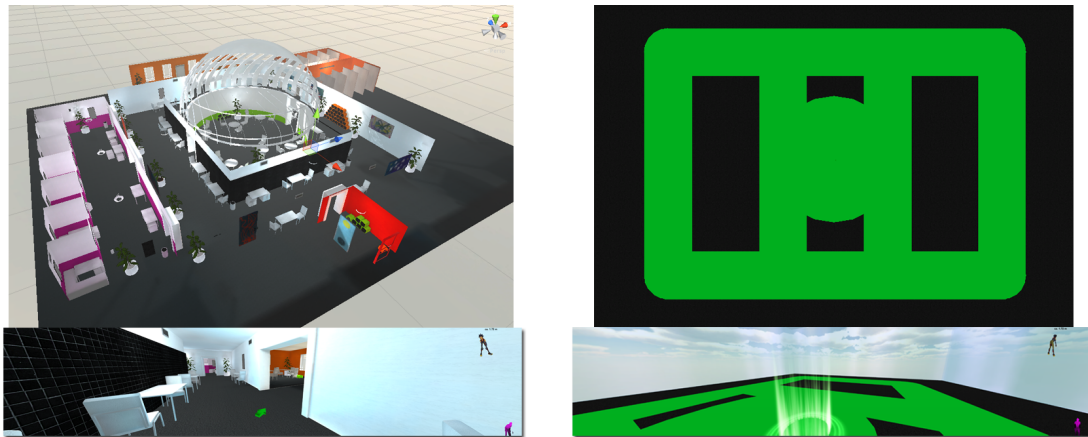


Figure 3.4.: KinectWalker. On the left: the training level. On the right: the level for the main task. Both shown from a bird's-eye and from a first-person view.

The second task was a memorizing task, where the participants were asked to memorize digits in the order in which they heard them. A digit was read out loud by the director of studies during the exercise whenever the participant collected a landmark. After he had collected all of the landmarks, the participant was asked to recall the digits in the right order and to write them down. The participant started the main exercise by performing the same waving gesture like he did in the training exercise.

After three iterations, each consisting of one training and one main exercise, the post test questionnaire (see appendix A.4) was handed to the participant. After that, the user study was finished.

### 3.2.3. Questionnaires

During the study, a participant was asked to fill out three different kinds of questionnaires, a pre test questionnaire, a post test questionnaire and a questionnaire to rate the perceived cognitive workload for the three tested interaction models.

#### Pre Test Questionnaire

The pre test questionnaire (see appendix A.1) gathered demographic data from the participant. The questions were (original questions were written in German):

- *How old are you?*  
Answer could be given in pre-defined ranges.
- *Did you play a computer game in the last four years?*  
The answer was *yes* or *no*. This question should show if the participant is used to modern VR software.
- *Which kind of computer game did you play?*  
Multiple pre-defined answers and a field for optional other answers were given. The question should describe in more detail to which kind of VR software the participant might be used to.
- *Do you have a Kinect sensor?*  
The answer could be *yes* or *no* and should show if the participant is used to VR systems using the Kinect sensor.
- *Do you use your computer for learning?*  
Possible answers were *yes* or *no*.
- *How do you use your computer for learning?*  
Multiple pre-defined and an optional free text answer were possible. In combination with the before mentioned question, this information showed in how far a participant is used to eLearning software.
- *How long do you usually use your computer per day?*  
Single answer out of a pre-defined list could be chosen. The question should show how experienced a participant is in using a computer.

#### NASA TLX

Since each of the selected interaction models should be rated by the participants in terms of perceived cognitive load, a questionnaire was handed to them directly after each main exercise. That means that the participants had to fill out this questionnaire three times.

As rating, the NASA Task Load Index (TLX) (Hart & Staveland, 1988) was selected because it is widely used in the field of HCI research (e.g., Battiste & Bortolussi, 1988; Byers, Bittner, Hill, Zaklad, & Christ, 1988)) and its structure fits well into what should be examined in this thesis.

The NASA TLX is a “multi-dimensional rating scale [...], in which information about the magnitude and sources of six workload-related factors are combined to derive a sensitive and reliable estimate of workload” (Hart & Staveland, 1988) in order to assess a task. Thus, the NASA TLX describes a tool to measure the perceived cognitive workload of a user who has performed a certain task. The index was developed by the Human Performance Group at NASA’s Ames Research Center in 1986. Compared to other tools like the Subjective Workload Assessment Technique (Reid & Nygren, 1988) or the Workload Profile (Rubio, Díaz, Martín, & Puente, 2004) the NASA TLX provides the best criteria to be used within this study. It is easy to be administrated as paper and pencil version and according to Rubio et al. (2004) it shows a high correlation with performance. Since the subjective results from the user study of this work should also be compared to the objective measured performance data of the main tasks, the NASA TLX was chosen as evaluation tool.

The questionnaire is divided into two parts. In the first part (see appendix A.3.1), the participant has to rate six dimensions, where each one describes a factor for the overall perceived workload within a numeric scale between 1 and 20. The values are later multiplied by 5 to rate each dimension within a 100-point range with 5-point steps. The six dimensions are further explained to the participant by an additional descriptive question. According to Hart and Staveland (2002), these questions are:

- **Mental Demand:** How mentally demanding was the task?
- **Physical Demand:** How physically demanding was the task?
- **Temporal Demand:** How hurried or rushed was the pace of the task?
- **Performance:** How successful were you in accomplishing what you were asked to do?
- **Effort:** How hard did you have to work to accomplish your level of performance?
- **Frustration:** How insecure, discouraged, irritated, stressed, and annoyed were you?

In the second part of the questionnaire (see appendix A.3.2), the participant has to compare each of the above mentioned six dimensions pairwise with each other. For each of the 15 pairs he has to choose one of the two dimensions, selecting the dimension that

was more relevant to the cognitive workload he perceived. The result of this second part of the questionnaire describes a weighted factor for each of the six dimensions. This weight is then combined with the results from the first part to create the final NASA TLX for a task.

### **Post Test Questionnaire**

After the participant had completed the practical part of the study, the post test questionnaire was handed to him (see appendix A.4). This questionnaire was used to get feedback to anything the participant might found good or bad about the study and to get recommendations for improving the presented interaction models. In addition, two simple questions were asked, namely which of the interaction models was the best and which one was the worst according to the participant's opinion.

- *Which interaction model is the best in your opinion?*  
The answer could be *Leaning*, *Running* and *Circle*, which indicate the three interaction models LEAN, WiP and D2V.
- *Which interaction model is the worst in your opinion?*  
The possible answers were the same as in the before-mentioned question.
- *Do you have any improvement suggestions to give or are there any aspects of the interaction models which you can think of in different combinations?*  
The answer could be given as free text.

### 3.3. Results

This section will present and evaluate the most important findings from the conducted user study. A detailed overview of the raw data, covering the answers from the NASA TLX questionnaires as well as the automatic measured performance data, can be found in appendix A.5. The study took place between mid of July and beginning of August 2014. The last participant took part in the study on 2 August 2014.

#### 3.3.1. Demographic Data

22 participants took part in the user study, out of which 8 participants were between the ages of 10 and 14. 13 participants were between 21 and 40 years old and 1 participant was older than 40. Overall, 7 participants were female and 15 male. Out of the female participants, only one was between the ages of 10 and 14. The other 6 female participants were between 21 and 40. Only 4 of all participants did not play any computer game in the last 4 years. The average PC usage per day was about 2,95 hours, where one participant did not use his PC daily and some participants took their time at work into consideration when answering the question. For simplicity, the answer “more than 3 hours” ( $> 3$ ) was counted as 4 hours to compute the average of the daily PC usage.

Participants who own a Kinect sensor and/or played a computer game from the categories adventure, egoshooter, simulation or Wii (which one of the participants added as additional kind of game) within the last four years were classified as users with a high experience in VR (*VRhigh*). Those participants who do not belong to this group but played a game of another category in the last four years were classified as medium experienced in VR (*VRmid*). All other participants were classified as inexperienced in VR (*VRlow*).

#### 3.3.2. Perceived Workload

When taking into account only the results from the NASA TLX questionnaire, the average perceived workload of all participants shows that the interaction model D2V produced the lowest workload (TLX = 38,61), compared to the LEAN (TLX = 48,98) and WiP (TLX = 58,29) models (see fig. 3.5). Lower TLX values indicate that the according model provided more capacity of working memory for other cognitive tasks beside the navigation in the VR world. Unfortunately, there are no thresholds given in the literature that provide an explicit interpretation of the absolute value of the NASA TLX. Nevertheless, the values show that all three interaction models are located in the mid range of the index’s 100-point scale.

Having a further look at the results for each of the six dimensions from the first part

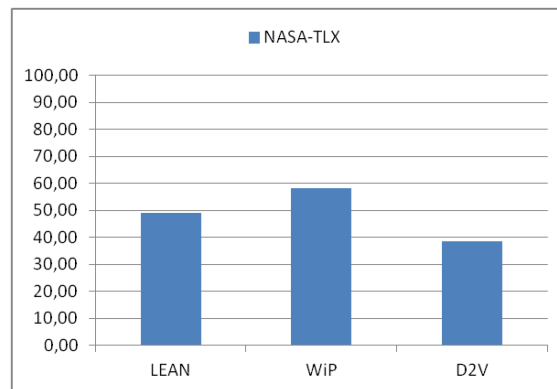


Figure 3.5.: Overall NASA TLX for each of the tested interaction models.

of the questionnaire, you can see that D2V outperformed WiP and LEAN in any of the six dimensions (see fig. 3.6). For all models, the dimension *Temporal Demand* shows a very low impact on the cognitive workload. This could be caused by the fact that most of the participants were mainly concentrating on the memorizing task rather than on completing the exercise as fast as possible, although, they were aware of the completion time being recorded as well. Some participants even waited with collecting the last landmark during the study in order to use the additional time for recalling the digits before they were asked to repeat them. This is also reflected in the results for the dimension *Mental Demand*, which, in contrast, show a high rating for all interaction models.

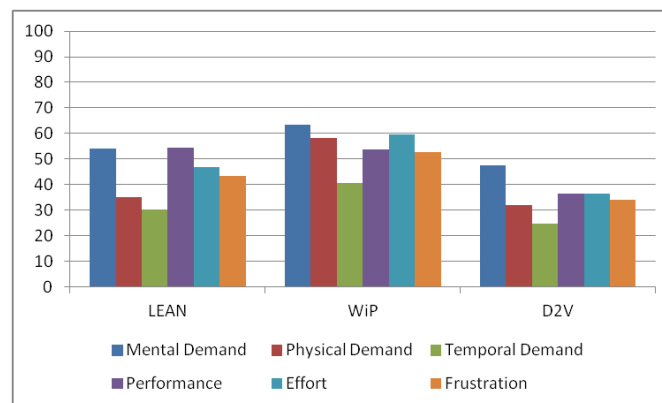


Figure 3.6.: Detailed results for each dimension and interaction model. A high value indicates a high impact on the perceived workload.



To evaluate the given answers from the post test questionnaire about the preferred interaction model, the following rating was used. If model  $x$  was marked as best by participant  $n$ , the model was rated with  $X_n = 1$ . If model  $x$  was marked as worst, it was rated with  $X_n = -1$ . If it was not marked to be the best or the worst model, it was rated with  $X_n = 0$ . The final score was then computed by the sum of  $X$  over all participants. Thus, the final score for a single interaction model  $x$  can be described by the simple equation

$$\text{Score}_x = \sum_{n=1}^{22} X_n.$$

Comparing the former results with this rating (see fig. 3.7), you can see that the results correlate with the evaluated perceived workload from the TLX (see fig. 3.5).

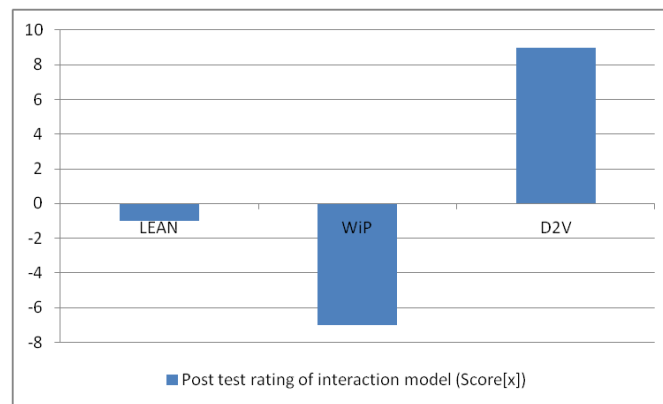


Figure 3.7.: Rating of interaction models according to the answers from the post test questionnaire about the best and worst model.

## Age

As described in section 3.3.1, 8 participants were aged between 10 and 14 years, 13 participants were aged between 21 and 40 and only 1 participant was older than 40 years. For the evaluation and comparison of the different age groups, the latter participant is removed from the results since he would otherwise define a whole age group as a single person.

Having a look at the overall NASA TLX of the interaction models depending on the factor *age* (see fig. 3.8), you can see that the results of both age groups show the same ranking of interaction models concerning the computed TLX as the results of all participants (cf. fig. 3.5). That means that D2V produced the lowest perceived cognitive workload, LEAN produced a medium amount of workload and WiP produced the highest workload.

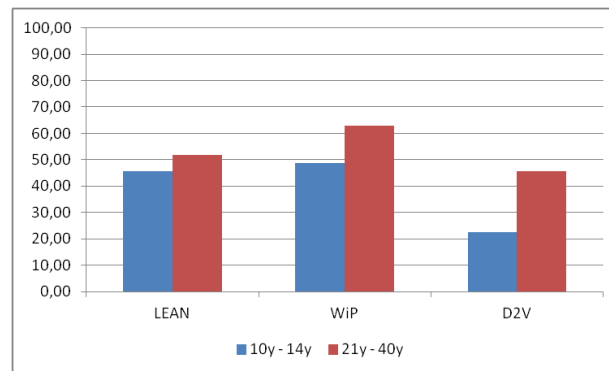


Figure 3.8.: Overall NASA Task Load Index for each of the tested interaction models depending on *age*.

What is striking when comparing the two age groups is that the young participants obviously perceived a much lower cognitive workload than the participants from the 21y-40y group when using the D2V model, while differences between the results for the WiP and LEAN model are not outstanding. For a statistical significance test of these values, the Mann-Whitney U-Test (McKnight & Najab, 2010) was used. This decision is based on the fact that the number of participants in the study is only 22 and a normal distribution of the measured data can therefore not be guaranteed. Thus, in order to ensure a statistically correct evaluation, the Mann-Whitney U-Test is preferred over a simple unpaired two sampled t-test.

The null hypothesis  $H_0$  claims that the factor *age* does not have an effect on the perceived workload, while the alternative hypothesis  $H_1$  claims that *age* influences the perceived cognitive workload of the participants. The test showed that for all models the results do not significantly differ between the groups of young and old participants ( $p_{LEAN} = 0.271$ ,  $p_{WiP} = 0.139$ ,  $p_{D2V} = 0.077$ ,  $\alpha = 5\%$ ). Thus,  $H_0$  cannot be rejected for any of the three models. However, since the difference for D2V barely misses the significance level of 5% in contrast to LEAN and WiP, it can be assumed that the factor *age* might have a certain effect on the perceived cognitive workload for the D2V model.

Comparing the six rated dimensions of the NASA TLX depending on *age* (see fig. 3.9), it can be shown that young participants seem to need less effort to complete the main task than older participants when using any of the three interaction models. The results also show that the dimension *mental demand* had ratings greater than 50 points in any model and in both groups except for the D2V model in the group of young participants. It was even rated with a low value of only 28,75 points in this case. This value is also reflected in the statements that some of the young participants made during the main task. They mentioned that since the avatar moves continuously forward when they in-

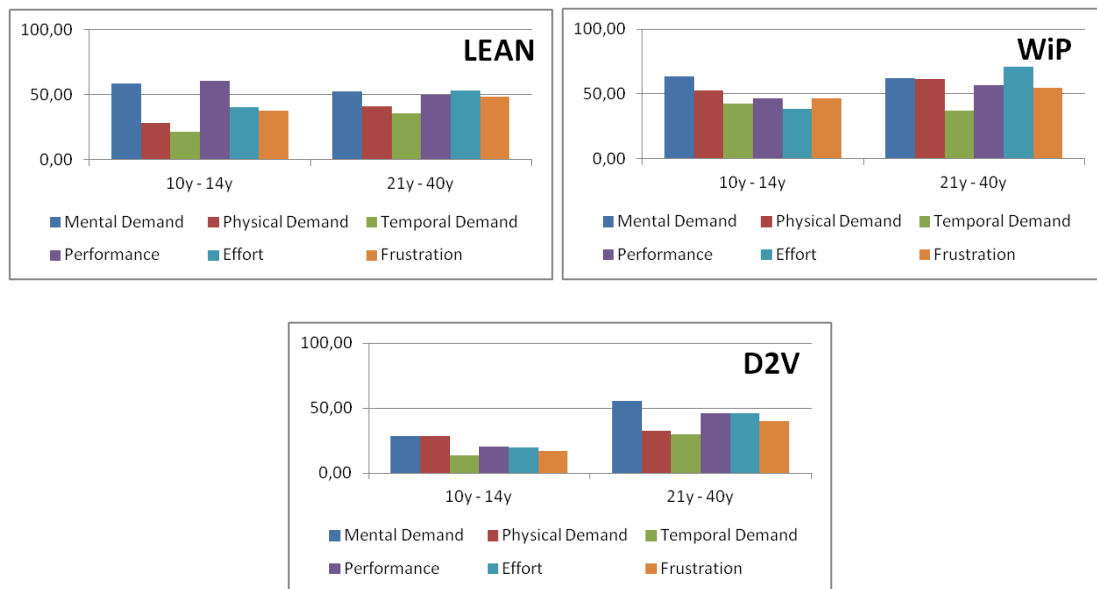


Figure 3.9.: Detailed results for each dimension and interaction model depending on the age of the participants. A high value indicates a high impact on the perceived workload.

crease the distance to the center point of the interaction area, they have more time to concentrate on the memory task because they do not have to hold a certain pose for the whole time. One should take into account that these observations do not correlate with the assumptions made in chapter 2 according to which D2V represents the model with the lowest abstraction from real life walking behavior. However, the results also correlate with the low rated *effort* dimension for the D2V model, especially in the group of young participants.

## Gender

The variable *gender* is another factor that should be investigated in detail since “men and women think and feel differently, and perceive, value and understand the world in their own ways” (Liu & Mihalcea, 2007).

As mentioned before, 7 participants were female and 15 participants were male. Again, the mean values of the NASA TLX for both groups show the same ranking of models as the overall average. In general, it can be seen that male participants perceived a slightly lower cognitive load than females for WiP and D2V, while the results for the LEAN model are nearly identical. In the same way as it was tested for age groups, none of the models provided significant differences between both groups according to the Mann-

Whitney U-Test ( $p_{LEAN} = 0.834$ ,  $p_{WiP} = 0.204$ ,  $p_{D2V} = 0.159$ ,  $\alpha = 5\%$ ).

Nevertheless, while D2V outperformed the other models clearly in the former discussions, it reached nearly the same TLX value as the LEAN model for female participants ( $TLX_{LEAN}[female] = 50,05$  and  $TLX_{D2V}[female] = 49,62$ ). This is interesting because it indicates that female participants might have more difficulties in using the D2V model than male participants. However, we have to consider that there was only one young, female participant taking part in the study, who rated the D2V model as best model with a NASA TLX of 18 points. Thus, we must assume that the high mean value for D2V is mainly caused by older female participants. In how far the effect is correlated to *gender* rather than to *age* cannot be further evaluated with the results from this study due to the before-mentioned fact that there was only one young, female participant taking part in the study. Another possible reason could be found when evaluating the results depending on VR experience (see section 3.3.2) since the female participants that rated D2V worst do also belong to the group of low experience VR users.

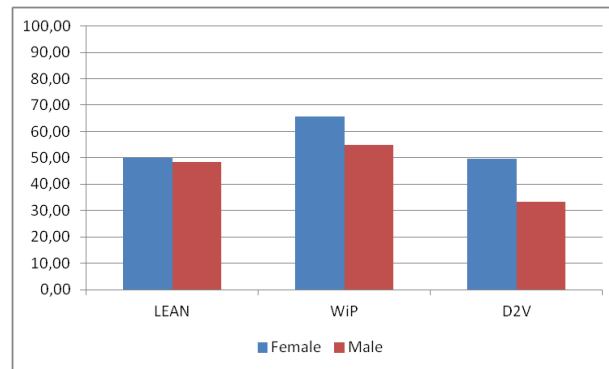


Figure 3.10.: Overall NASA Task Load Index for each of the tested interaction models depending on *gender*.

A further look at the detailed results of the six NASA TLX dimensions (see fig. 3.11) confirms the above-mentioned observations regarding the ranking of the models and it supports the assumption that male participants, in general, perceived a lower cognitive workload than female participants. The results also correlate with the overall results from fig. 3.5 since the dimension *mental demand* was rated high through all models and groups while the dimension *temporal demand* was rated low. Similar to the comparison of young and old participants, the dimension *mental demand* was rated much better for D2V in one group. Male participants rated the *mental demand* for the D2V model significantly lower than females ( $p = 0.031$ ,  $\alpha = 5\%$ ). This might indicate that female participants were much more challenged by the navigation task than male participants when using D2V. Thus, the results support the before-mentioned effect that D2V had on the perceived workload for D2V in the female group.

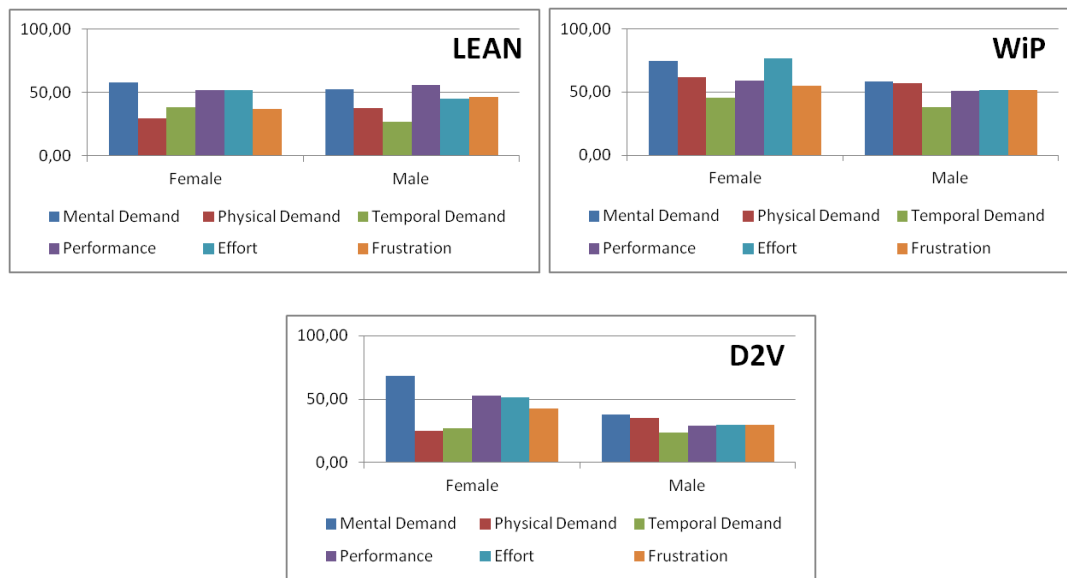


Figure 3.11.: Detailed results for each dimension and interaction model depending on the gender of the participants. A high value indicates a high impact on the perceived workload.

## VR Experience

In the same manner as it was done for *age* and *gender*, the results for the three different groups of participants *VRlow*, *VRmid* and *VRhigh*, as they were defined in the introduction of this section, were evaluated. The overall NASA TLX (see fig. 3.12) again correlates with the ranking from the overall mean results of fig. 3.5, concerning the two groups *VRhigh* and *VRmid*. What is striking, however, is the ranking of the models in the *VRlow* group. In this group, the LEAN model (TLX = 43,75) was the one preferred over D2V (TLX = 64,67) and WiP (TLX = 74,75). The models WiP and D2V were even rated worse than the worst model from the overall NASA TLX in fig. 3.5. Taking into account the results for the different dimensions of the NASA TLX (see fig. 3.13), it can be assumed that the perceived frustration in the group *VRlow* could be one criterion for the high ratings of the D2V (frustration: 75) and the WiP (frustration: 73,75) compared to the LEAN model (frustration: 27,50). The high level of perceived frustration might be caused by the fact that the participants in the *VRlow* group were not used to systems like the Kinect sensor and therefore had difficulties in using a HCI that provides natural interaction. As assumed in chapter 2, LEAN was defined as the model with the highest level of abstraction while the other two were said to be more natural regarding walking behavior in real life. In contrast to *VRlow*, the participants from the groups *VRmid* and *VRhigh* might also be aware of the fact that such systems most often provide only limited accuracy in terms of e.g. recognizing gestures. Since the models WiP and D2V

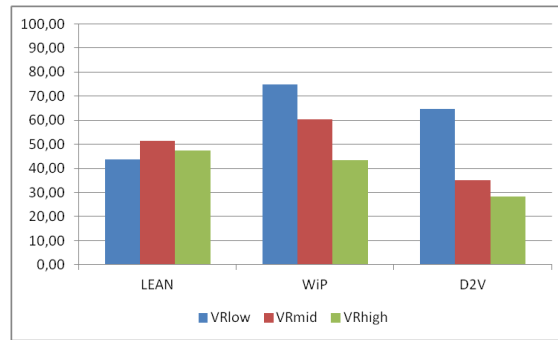


Figure 3.12.: Overall NASA Task Load Index for each of the tested interaction models depending on *VR experience*.

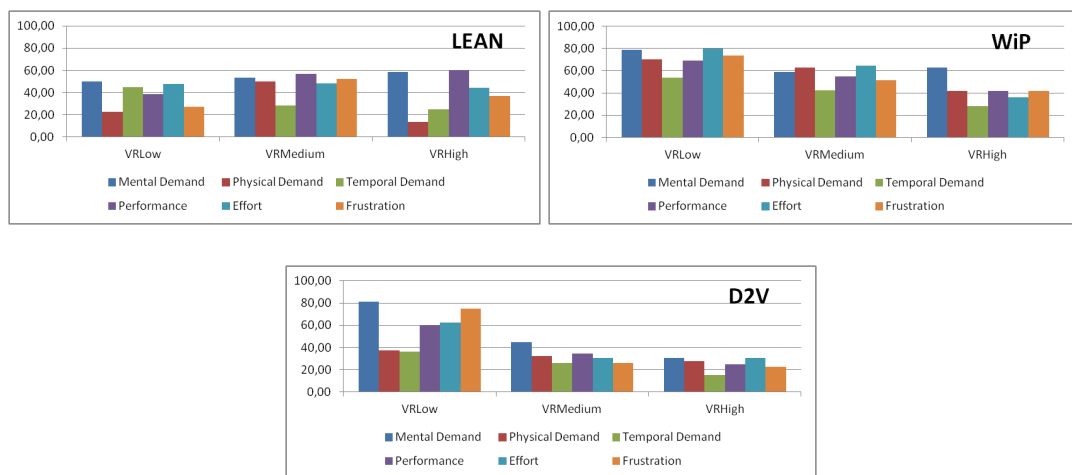


Figure 3.13.: Detailed results for each dimension and interaction model depending on the VR experience of the participants.

have been defined as being more natural than the LEAN model (cf. chapter 2), they are also based on a more complex implementation and, therefore, might be less accurate in detecting a user's pose, which could also have influenced the perceived frustration.

To extend the results, another comparison has been investigated. For this purpose, the participants were divided into the two groups *Kinect* (4 participants) and *No Kinect* (18 participants), referring to those participants that own a Kinect sensor and those who do not. The results of the average NASA TLX for those groups (see fig. 3.14) support the before-mentioned assumptions for the groups *VRlow*, *VRmid* and *VRhigh*. While LEAN is rated with a TLX between 45 and 50, the differences between WiP and D2V are much higher between the groups *Kinect* and *No Kinect*. Unfortunately, these observations cannot be confirmed by using the Mann-Whitney U-Test for significance since

the sample size of the *Kinect* group is too small. However, the results might support the before-mentioned assumption that participants with less experience in VR have more difficulties in using natural HCIs. These findings also show that models which are said to be natural do not have to be more intuitive than abstract models.

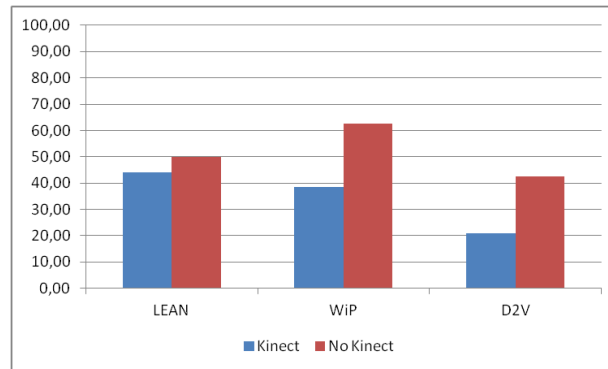


Figure 3.14.: Overall NASA Task Load Index for each of the tested interaction models depending on whether the participant owns a Kinect sensor (*Kinect*) or not (*No Kinect*).

### 3.3.3. Measured Performance

The following section will investigate the objective measured performance data from the user study and tries to link the findings to those from the subjective measured data of the last section. In the following, the term *error* will describe the state in which a participant left the pre-defined, green walk path of the main level. Each time he left the path, an error was counted. Furthermore, the time during which the participant was off the path was measured and used to get a relation between this time of error and the completion time of the task.

On average, each participant needed about 25 minutes to complete the study. The average time for completing the main task using LEAN was about 102 seconds, 101 seconds for WiP and 83 seconds for D2V (see fig. 3.15). Thus, D2V outperformed LEAN and WiP by more than 20 seconds on average. The average amount of errors that were made during the main task was 7,59 using LEAN, 4 using WiP and 5,5 using D2V (see fig. 3.16), while the time of error in relation to the completion time was about 22 % on average for each model.

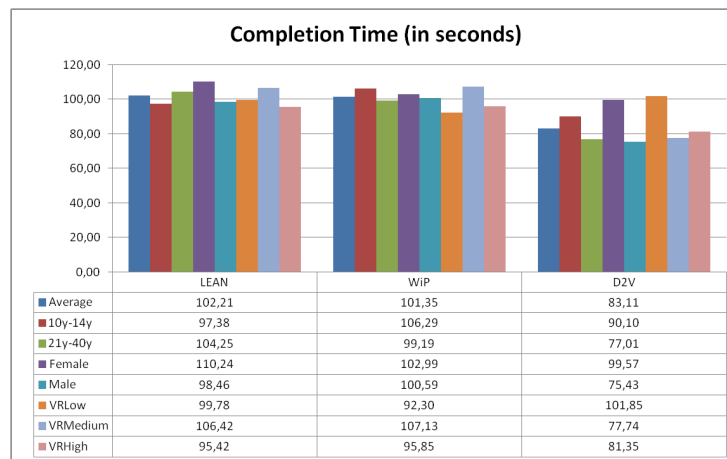


Figure 3.15.: Completion times of the main task.

Young participants aged 10 to 14 performed best when using the D2V model and worst when using the WiP model, considering only the completion time (difference between D2V and WiP: 16 seconds). However, they made only 3,38 errors on average when using WiP. These values correlate with the ranking of models from fig. 3.8 when taking the completion time into account, but mismatch the subjective ranking when considering the amount of errors as performance measurement. In comparison to the young participants, the group of participants aged 21 to 40 performed better when using WiP or D2V. When using D2V, the older participants even made less mistakes on average.



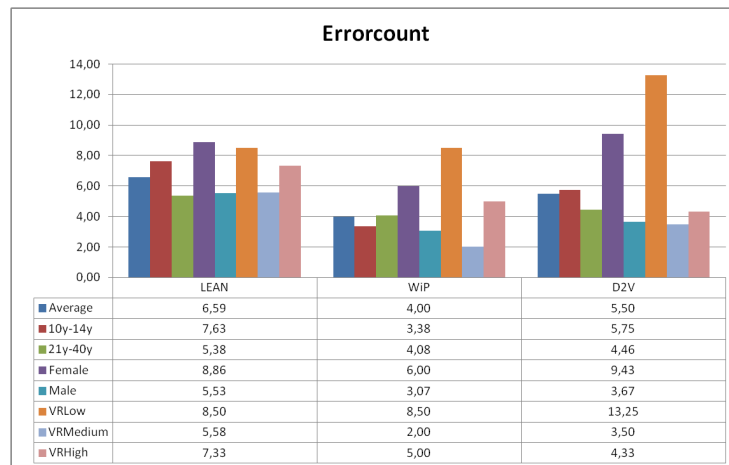


Figure 3.16.: Amount of errors that were made in the main task.

The performance results for this group also correlate with the subjective ranking of fig. 3.8 considering completion time as only measurement, and they mismatch with regard to the amount of errors.

In case of the variable *gender*, the performance results confirm the observation that male participants performed better than female participants. While the completion time of both groups differ only slightly ( $p_{LEAN} = 0.204$ ,  $p_{WiP} = 0.944$ ,  $p_{D2V} = 0.180$ ,  $\alpha = 5\%$ ), the differences between the amount of errors are noticeably high between the two groups ( $p_{LEAN} = 0.194$ ,  $p_{WiP} = 0.342$ ,  $p_{D2V} = 0.061$ ,  $\alpha = 5\%$ ). Again, the results correlate with the according ranking from the subjective evaluation in the group of male participants when considering completion time as performance measurement and they mismatch for the amount of errors. WiP reached the best values in terms of errors being made for both male and female participants. For females, the objective measured data mismatches with the subjective data regarding completion time as well as the amount of errors. An interesting observation is also the relation between completion time and time of error for female participants since they were in the state of error for more than 28% of the time (see fig. 3.17) when using LEAN and even more than 38% when using D2V. This confirms the assumption that female participants had more difficulties in the VR system than male participants.

Finally, for the three groups of participants with different levels of VR experience, it can be observed that they performed almost the same when using LEAN and WiP, but *VRmid* and *VRhigh* outperformed *VRLow* for D2V in terms of completion time (*VRLow*: 102 seconds; *VRmid*: 78 seconds; *VRhigh*: 81 seconds). However, *VRhigh* performed slightly worse than *VRmid* when using D2V, which was not expected from the subjec-



Figure 3.17.: Time of error that was measured in the main task.

tive measured data. It is important to notice that *VRmid* outperformed both other groups in terms of counted errors. Using the WiP model, *VRmid* achieved an average amount of errors of only 2 while the completion time was the worst compared to the other models in this group. Another very interesting result is the time of error for the *VRlow* group. Participants of this group resided almost half of the time in the state of error when using WiP or D2V and 29% of the time when using LEAN. This confirms the assumption of participants from the group *VRlow* having difficulties in using VR systems and natural HCIs.

### Memorization Task

This section investigates the measured performance from the memorizing task of recalling numbers in the right order during the main task. The number of correct digits was measured in two different ways. The first measurement is based on the number of correct digits from the beginning of the sequence until the first mismatch is detected. This method is named *correct digits in prefix* (see fig. 3.18). The second method is called *correct digits at position* and compares the digits of the original and the recorded number per arity, i.e. digits are compared at each index. While the first method uses a stronger criterion in terms of correctness, the second one was used to show the more general memorizing capabilities of the participants. In general, the results for the second method are approximately one digit higher than for the first method.

It can be observed from fig. 3.18 and 3.19 that D2V performed best for the method *correct digits in prefix* (4,14 digits in prefix on average), but WiP performed best for the method *correct digits at position*, for which D2V performed even worst. It is striking

that the group *VRlow* reached very high values when using LEAN for both methods (6,5 digits in prefix and 7,25 digits at position on average). This correlates with the ranking of models concerning the perceived workload for the *VRlow*, where LEAN was rated to have a lower workload than the other two models. It is also striking that for the other models, the performance of the group *VRlow* was drastically worse. This might, again, indicate the difficulties that untrained and VR inexperienced users have when they are confronted with natural interaction techniques rather than abstract input commands.

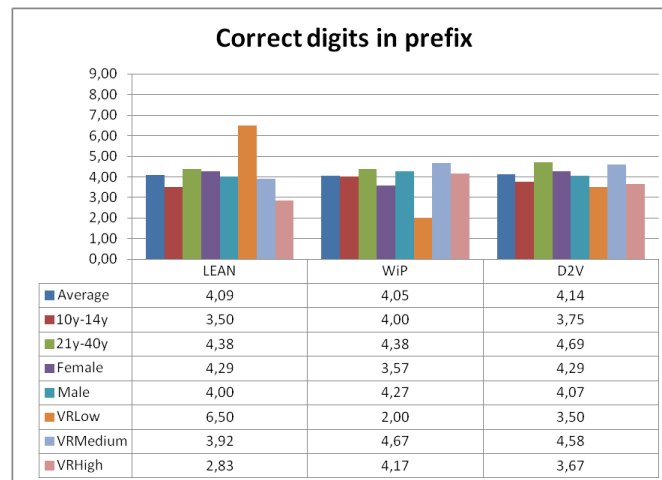


Figure 3.18.: Results of the memorizing task using the method *correct digits in prefix*.

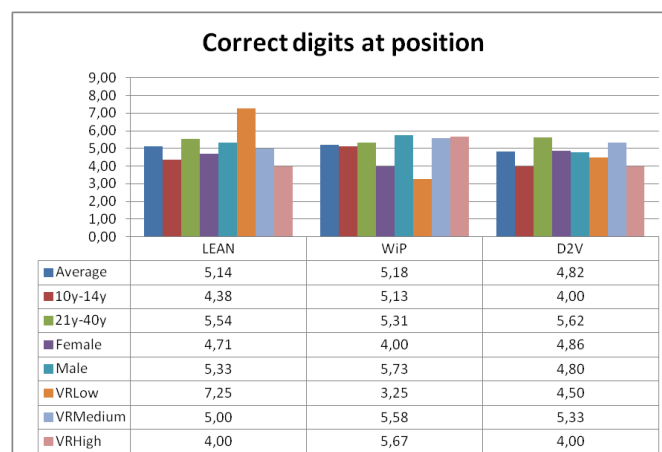


Figure 3.19.: Results of the memorizing task using the method *correct digits at position*.

## Navigation Task

At this point some of the recorded walk paths of the participants will be presented in order to confirm some of the before-mentioned assumptions about the different groups of participants. The landmarks originally appeared in the order that is visualized by fig. 3.20. The image also shows the intended walk path to collect all landmarks. An overview of all recorded walk paths can be found in appendix A.6.

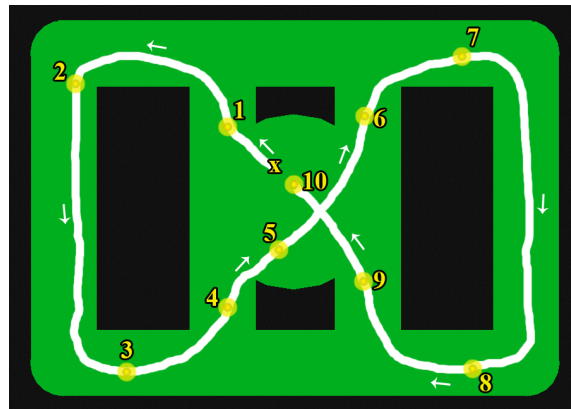


Figure 3.20.: Intended walk path of the main task. The participants started at position  $x$  and should collect the landmarks in ascending order.

As a first example, the walk path of participant 3 (10y - 14y, male, VRhigh) is given in fig. 3.21. In contrast to this, fig. 3.22 shows the walk path of the adult participant 14 (21y - 40y, male, VRmid). What can be observed for both participants is that they made most of their errors during the last part of the task. While the first walk path (LEAN) of both participants shows that landmarks were missed several times, it can be assumed that collecting those landmarks was easier when using WiP or D2V. However, the paths do not reflect the large difference between the NASA TLX ratings of the D2V model for young and adult participants.

The next example shows a representing walk path from the group of female and VR inexperienced participants (see fig. 3.23), whereas fig. 3.22 is an example for male and VR experienced participants. It seems obvious that the female participant had much more difficulties in completing the navigation task than the young, male participant. Especially the walk paths for D2V indicate a high cognitive workload for females, which correlates with the high rated mental demand for D2V in the female group. It also confirms the observation that VR inexperienced participants performed best when using an abstract interaction model like LEAN.

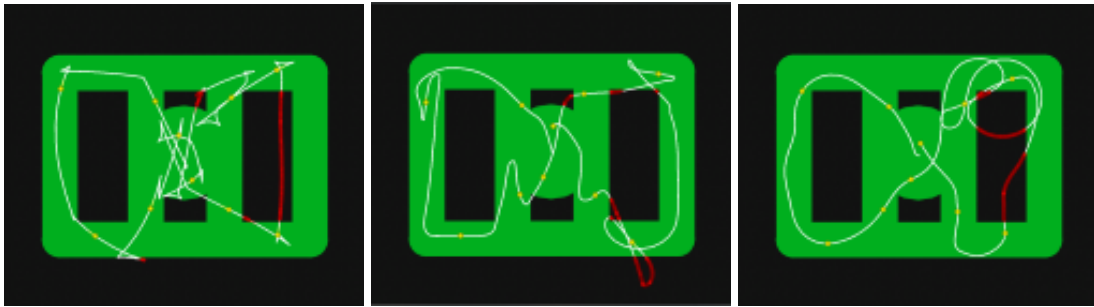


Figure 3.21.: Walk paths of participant 3 (10y - 14y, male, VRhigh). From left to right: LEAN, WiP, D2V. Red lines indicate positions off the path while white lines indicate positions on the path. Yellow dots are landmarks.

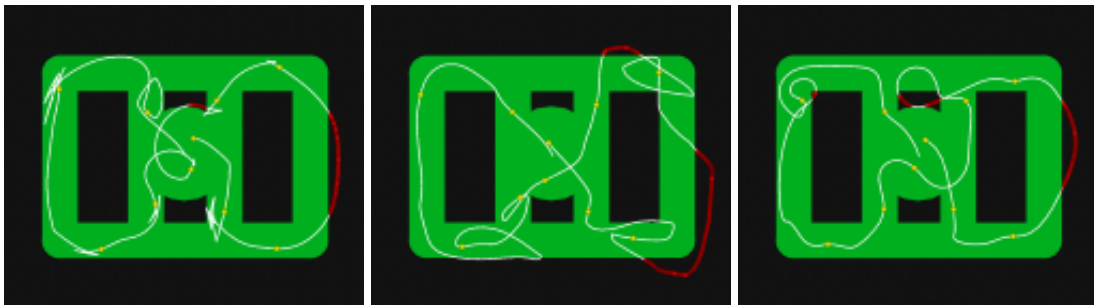


Figure 3.22.: Walk paths of participant 14 (21y - 40y, male, VRmid). From left to right: LEAN, WiP, D2V.

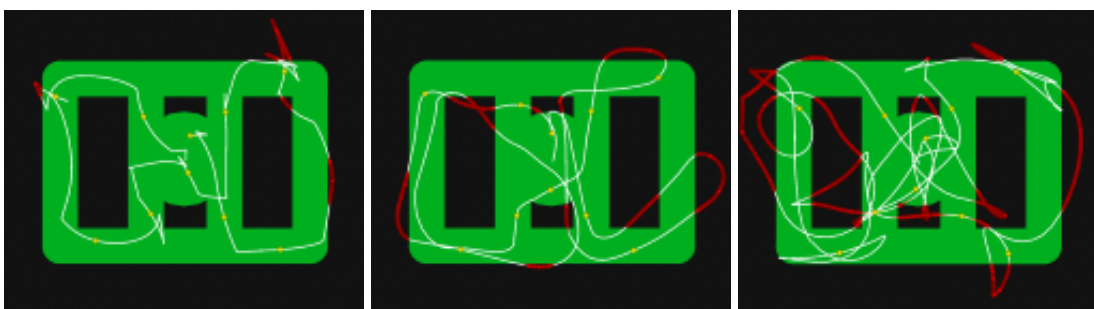


Figure 3.23.: Walk paths of participant 21 (21y - 40y, female, VRlow). From left to right: LEAN, WiP, D2V.

### 3.3.4. Feedback

The evaluation of the free text answers from the post test questionnaire reveals further subjective information about the training setup and the interaction models, but it also provides recommendations of how the models might be improved for further studies.

For the training setup, some rather irrelevant feedback was given, like improving the arrow or making the training scenario more interesting by adding music and more action. Such comments were not taken into consideration for further investigations. Some participants came up with useful ideas for the further development of the selected models. For example, several people mentioned that the missing possibility of going backwards is counterproductive for the WiP model. For others, it was hard to break when using D2V, especially when they lost their orientation towards the center point. Participant 10 even had problems to stop movements when using LEAN.

As recommendations for improving the models, three of the participants mentioned that combining the D2V model with the features of LEAN could be an improvement. That would cover a leaning gesture for accelerating and spatial movements for indicating the direction. One participant even suggested to replace the implemented rotation gesture with a leaning gesture. Another suggestion was to combine the D2V model with the WiP model, which means that WiP is used for acceleration and D2V for indicating the direction. Both recommendations are valuable and should be considered in possible future studies using the *KinectWalker* application.

### 3.3.5. Conclusion

To sum up the subjective measured data from the study, the results showed that the three interaction models all achieved a medium or even better score in the NASA TLX scale ( $TLX_{LEAN} = 48, 98$ ,  $TLX_{WiP} = 58, 29$ ,  $TLX_{D2V} = 38, 61$ ). The model with the lowest perceived cognitive workload according to these results is the D2V model, the second best model is LEAN and the worst of these models is WiP. The same result could be computed when considering the ratings for the best and the worst model in the post test questionnaire. This ranking also holds for groups of different gender or ages as well as for medium and high experienced VR users. Only those participants that were said to have a low VR experience prefer the LEAN model over D2V and WiP. While the differences are not significant, we can assume young participants to have perceived a measurable lower cognitive workload than older participants, especially when using D2V ( $p_{LEAN} = 0.271$ ,  $p_{WiP} = 0.139$ ,  $p_{D2V} = 0.077$ ,  $\alpha = 5\%$ ). Thus, it can be assumed that the factor *age* might have a certain impact on the workload in combination with the D2V interaction model. In addition, it was shown for male participants that they perceived a lower cognitive workload than female participants for each of the selected

models. However, the differences were not significant ( $p_{LEAN} = 0.834$ ,  $p_{WiP} = 0.204$ ,  $p_{D2V} = 0.159$ ,  $\alpha = 5\%$ ). Recapitulating hypothesis  $H_2$  (*The demographic parameters gender and age of a participant have a significant effect on the perceived workload.*), it can be concluded that there is no significant impact of both variables in general, whereas the variable *age* showed a high impact when being investigated for only the D2V model.

Concerning  $H_1$  (*The level of abstraction of an interaction model compared to real walking behavior affects the perceived cognitive load of the participants. The more natural a model is, the lower the workload.*), it can be summarized that D2V, being defined as the most natural interaction model according to chapter 2, outperformed LEAN and WiP in terms of perceived workload in general. Nevertheless, WiP, as second most natural model, performed worst on average while LEAN was rated to provide the lowest cognitive workload for VR inexperienced participants in contrast to the groups *VRmid* and *VRhigh*. Thus,  $H_1$  has to be rejected.

Hypothesis  $H_3$  (*Participants without or with less experience in VR will show a lower cognitive performance than participants with VR experience.*) must also be rejected since the results from the memorizing task reveal that participants from the group *VRlow* even outperformed *VRmid* and *VRhigh* when using the LEAN model. In addition, *VRmid* performed better than or similar to *VRhigh* for all three models.

Finally, the validation of hypothesis  $H_4$  (*The model that provides the best results in the memorizing task also outperforms the others in the training scenario considering completion time and time of error.*) depends on the method that is used for evaluating the memorizing task. If the method *correct digits in prefix* is used, D2V is the best model, which, indeed, outperformed the other models in terms of completion time and time of error. Nevertheless, if the method *correct digits at position* is used, the best model is WiP, which only outperformed the other two models when regarding the amount of errors being made.

The study could not confirm any of the assumptions concerning the cognitive load when using different interaction models in the presented VR setup. Nevertheless, the findings indicate that the parameters *age* and *VR experience* might affect the perceived cognitive workload. Since the measured results did not show significant differences in general for the small sample of participants that took part in this user study, further studies might confirm or reject the observations by concentrating on those parameters and acquiring more participants.

As a guideline for selecting a good interaction model for an effective VR learning application, which is based on gesture-based locomotion interaction, it could be a suggestion

to choose D2V in general and switch to one of the other models if detailed demographic data from the users is given. Thus, if a user is e.g. inexperienced in VR, one could prefer LEAN over WiP and D2V.



## 4. Application - KinectWalker

This chapter will describe in detail the concept, the design and the implementation of the *KinectWalker* application, which has been used to conduct the user study from chapter 3. The end of this chapter will describe how *KinectWalker* was integrated into the SafeChild project.

### 4.1. Concept

*KinectWalker* is used in several ways. Firstly, the application can serve as a means of performing a user study, and secondly, it can easily be integrated into any other VR applications to enrich those with a locomotion-based interaction. Therefore, a modular and easy-to-maintain design of the application is essential.

#### 4.1.1. Requirements

The main feature of *KinectWalker* is the definition and recognition of different gesture-based interaction models to integrate locomotion behavior into any Unity3D application using the Microsoft Kinect sensor as the only physical input device. In addition, it should identify and log predefined actions within a VR environment during runtime, like a user's location in the virtual world when an issue regarding the training context is identified or the completion times of predefined tasks. *KinectWalker* is a VR application module, which has been developed to conduct the user study of this thesis. However, since the outcome of the user study should also lead to a prototype implementation as part of the SafeChild project, *KinectWalker* must meet several requirements, which influence its design and implementation.

#### Input Device

*KinectWalker* must be able to recognize a person in front of a workstation and further identify body points in order to define and recognize full body gestures. In detail, it must recognize, calibrate and track a user as well as provide positional and rotational information about the tracked body points. For this purpose, the Microsoft Kinect sensor was chosen as input device since it provides the needed functionality and is a low-cost

3D tracking device with a well-documented application programming interface (API) by Microsoft.

### **Framework for the User Study**

As *KinectWalker* is also intended to be used for the user study of this thesis, it must be able to administrate and control a pre-defined test scenario within a VR environment, i.e. defining and monitoring locomotion tasks in the test scenario, controlling a test session, logging certain actions that are relevant for later diagnosis of the test and also visualizing recorded data in a meaningful way.

### **Independent Toolkit**

The fact that the results of the user study in this thesis should also lead to a prototype implementation in the SafeChild project defines another important requirement. *KinectWalker* should be designed in such a way that it can be integrated into other projects easily. As for that, the overall design of *KinectWalker* ought to describe an independent toolkit, which can be adopted by other VR developers with little effort. Therefore, *KinectWalker* must be extendable by additional interaction models easily and it must be clear to other developers which parts of the application they have to adjust to address their needs.

## **4.2. Design**

This section will describe the abstract design of *KinectWalker* according to the requirements from section 4.1.1. It will clarify how *KinectWalker* approaches these requirements by design. In the first part, an overview of the different modules of *KinectWalker* will be given, showing how they are integrated and connected to each other. After that, a use case diagram will be shown to demonstrate how different roles of users can interact with the *KinectWalker* module.

### **General Design**

*KinectWalker* is not an executable application on its own. It rather defines a bundle of scripts, which is integrated into a surrounding VR system. Nevertheless, it acts as a single module when being integrated into other projects. Figure 4.1 shows an overview of how all the different modules in a VR system are connected when using *KinectWalker*. In the diagram, you can see the Kinect sensor on the right side. The *Kinect for Windows Software Development Kit (SDK)* is used to handle the hardware-based communication with the device. The processed data from the official Microsoft SDK is then wrapped

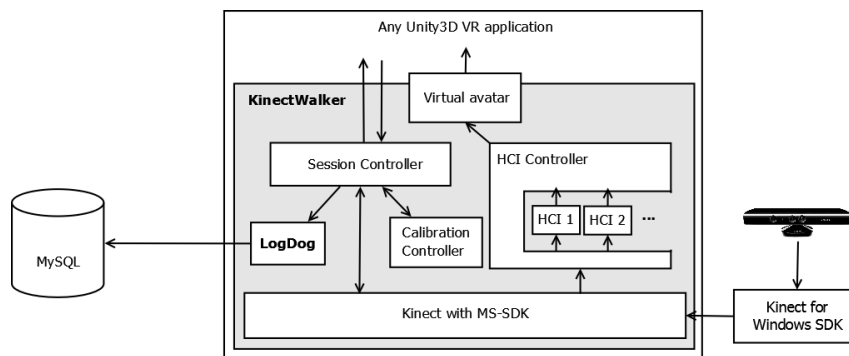


Figure 4.1.: Overview of the *KinectWalker* design within a VR application.

by the *Kinect with MS-SDK*<sup>1</sup> toolkit by *RF Solutions*, which is encapsulated into the *KinectWalker*, while *KinectWalker* itself is encapsulated into a surrounding VR application.

The core modules of the *KinectWalker* are the *Session Controller*, the *Calibration Controller*, the *HCI Controller*, the *LogDog* component and a *virtual avatar*.

1. **Session Controller:** The Session Controller is used by the surrounding VR application to administrate the locomotion behavior in the given virtual environment. If a test scenario is defined like it was done in the user study of this thesis, the Session Controller automatically switches between tasks, interaction models and exercises. It also allows the application to start and stop the tracking of users, i.e. enable or disable movements of the virtual avatar. Another function of the Session Controller is to inform the logging component of relevant actions from the training session that should be stored on the database, like the completion time of tasks or the time of error.
2. **Calibration Controller:** The Calibration Controller handles the identification of users and starts the process of calibration when a user takes up a certain calibration pose. It also measures the approximate body height of the user during calibration and adapts the height of the virtual camera in the 3D world accordingly. The latter function is based on the fact that adapting the camera's height should achieve a realistic view angle for users since they can perceive the surrounding environment from a view angle that is similar to their real life experience.
3. **LogDog:** The LogDog component is a simple tool that stores information on a connected MySQL database. Within the Unity3D Integrated Development Environment (IDE), it can also visualize stored information to the user. A more

<sup>1</sup><https://www.assetstore.unity3d.com/en/#!/content/7747> (last access 20 August 2014)

detailed description of the technical implementation of this module is given in section 4.3.3.

4. **HCI Controller:** The HCI controller is a controller class that routes the incoming data from the *Kinect with MS-SDK* tool to the available interaction models. It defines the active interaction model during runtime and gives VR developers the opportunity to add or remove additional models to or from their applications. If gestures of the active interaction model are recognized, the information about changes in rotation and translation is directly mapped to the virtual avatar.
5. **Virtual avatar:** The virtual avatar is a default *Character Controller* object in Unity3D, which can be moved around in a virtual 3D world.

## Use Cases

As mentioned in the last section, *KinectWalker* is not an executable application. However, it is a tool that users can interact with when it is integrated into a VR application. Therefore, the term *KinectWalker* is used as the name for an abstract component in the following use case diagram (see fig. 4.2). There are three roles that can interact with *KinectWalker*: **User**, **VR developer** and **director of studies**.

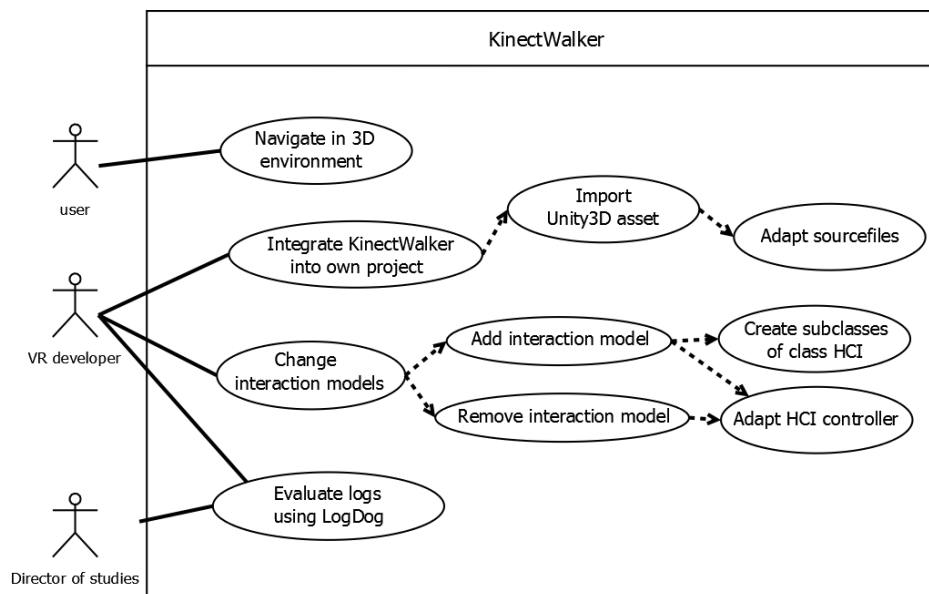


Figure 4.2.: Use case diagram of KinectWalker.

1. **User:** The user of *KinectWalker* - or the user of a system that integrates *KinectWalker* - is able to navigate freely through a virtual 3D world by using one of the available and activated interaction models of *KinectWalker*.
2. **VR developer:** A VR developer can integrate the *KinectWalker* module into his own project. To do so, he must import the bundled Unity3D asset, which installs the base version of *KinectWalker* into his Unity3D project. A Unity3D asset is a single file, which contains a bundle of scripts and other files. After the installation, the developer must adapt some of the imported files to meet his own needs. That is for example configuring the *HCI Controller* by adding or removing interaction models and defining one of them as active interaction model for his application. He can also change several visibility options, which e.g. show or hide the depth-map and RGB color-map of the recorded Kinect sensor information. Finally, if he is also the designer of a test scenario concerning locomotion behavior, he must also adapt the *Session Controller* to use it as an interface between the view layer and the input controller within his project.
3. **Director of studies:** The director of studies as well as the VR developer can use *KinectWalker* to evaluate recorded data directly within the Unity3D IDE. For that, he simply uses the *LogDog* toolkit that comes with *KinectWalker*. Further information about the technical details of the *LogDog* module is given in section 4.3.3.

## 4.3. Implementation

This section will describe the technical implementation of the *KinectWalker* application. The first part will introduce the Microsoft Kinect sensor in detail. The second part will list the software products that are used for the implementation. After that, the technical implementation of *KinectWalker* and the *LogDog* module will be presented. Finally, the last part of this section will describe how *KinectWalker* was integrated into the existing VR application SafeChild.

### 4.3.1. Microsoft Kinect Sensor

The Microsoft Kinect sensor is an optical tracking system and can be used as 3D input device for NUI-systems. Users interact marker-less and non-intrusive with such a system by using full body gestures and speech. The sensor was published in November 2010 by Microsoft as part of their gaming console XBox 360. Since then, it has evolved into a device that is used by a wide range of developers in the industrial and research area.

The Kinect sensor as it was used in this thesis is based on the reference design “The PrimeSensor” from 2010 (PrimeSense, 2010) and defines a combination of audio and visual sensors. The sensors are a RGB-camera, an infrared (IR) module, consisting of a IR-camera and a IR-emitter, and finally an array of four microphones (cf. fig. 4.3). In addition, a motor is attached to control the vertical view angle of the device.

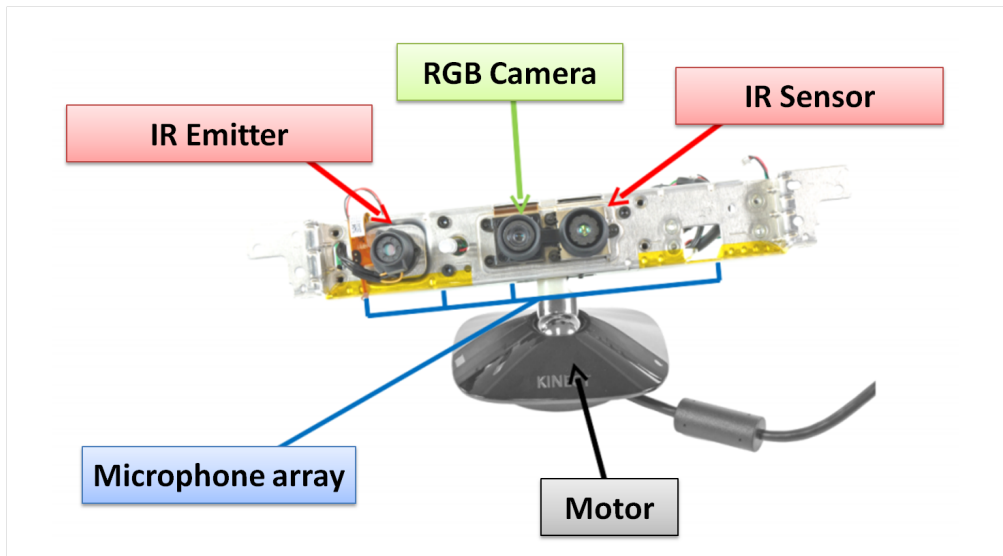


Figure 4.3.: The architecture of the Microsoft Kinect sensor. Modified after (Ifixit.com, 2010)

- **RGB-camera:** Takes pictures in a VGA resolution of 640x480 pixels with a color depth of 8 bit.
- **IR module:** The combination of IR emitter and IR camera works as depth sensor and takes depth pictures in a resolution of up to 640x480 pixels with a color depth of 11 bit.
- **Array of microphone:** The array of 4 microphones allows to locate audio sources in a room and is used to recognize voice commands from distinguishable users.
- **Motor:** The motor allows a physical change of the vertical view angle from -27 to +27 degrees.

The cameras of the Kinect sensor support a horizontal view angle of 57.5 degrees and a vertical view angle of 43.5 degrees with -27 to +27 degree tilt range up and down. Considering the official product description, the depth range in which the Kinect sensor operates is between 0.8 m and 4 m with noisier recognition when the range is extended

beyond 4 m. The sweet spot lays between 1.2 m and 3.5 m (Microsoft, 2013). However, Viager (2011) investigated these values and defined the sweet spot between 0.8 m and 3.5 m. He also confirmed the spatial resolution on the x-y-axis of 3 mm and a spatial resolution of 1 cm on the z-axis. (Viager, 2011)

In this thesis, the Kinect sensor of the first generation, which includes the above-mentioned technical features, was used. There are currently two more versions of the Kinect sensor available on the market. First, there is the *Microsoft Kinect for Windows*, which is a modified Kinect sensor of the first generation and which can only be used with computers. It introduced the so-called *Near Mode*, which decreases the minimum distance of the user to 40 cm. Secondly, there is the newest *Microsoft Kinect 2* sensor, which extends the former features with higher resolution of processed images (1080p - full HD) and wider vertical and horizontal view angles.

For all versions of sensors, the captured data is processed by the used SDK, which creates higher level information for a software application that is based on the Kinect sensor as its physical input device. The following section will list the SDK being used and other software products the implementation is based on.

### 4.3.2. Used Software

The *KinectWalker* is based on several third-party software products, which cover 3D models for the training scenario, an SDK, which is used to make sensor information available within the application, an IDE, in which the application is developed, and several tools that support the overall user experience of *KinectWalker*.

#### Kinect for Windows SDK v1.8

The *Kinect for Windows SDK* by Microsoft is used to communicate with the attached Kinect sensor in order to create applications that support gesture and voice recognition. The SDK consists of the drivers for the Kinect sensor, application programming interfaces that provide access to higher level information about captured data, device interfaces, and several tools to support the development.

The most important feature of the SDK in the context of this thesis is the possibility of tracking up to 20 body joints by providing positional and rotational information for each of them. Figure 4.4 shows these joints. Further features like face tracking and speech recognition were not used for the implementation of this work.

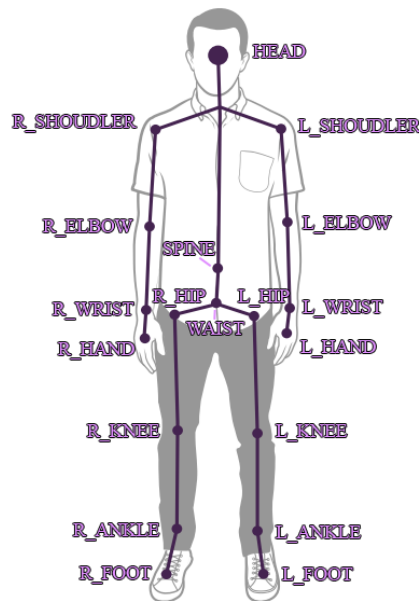


Figure 4.4.: The 20 body joints which can be tracked by the Kinect sensor (modified after (Microsoft, 2013))

## Unity3D

Unity3D is a widely-used game engine and IDE for 3D games and applications. It was published by Unity Technology and is used by more than 3 billion<sup>2</sup> developers worldwide. In comparison to similar game engines like e.g. the Unreal game engine, Unity3D is based on a freemium license model. That means that basic functionality is available for free, which, for the purpose of this thesis, covers all requirements regarding the development of *KinectWalker*. In addition, Unity3D is used for the development of the SafeChild project, which makes it easy to integrate *KinectWalker* into it.

A Unity3D project is defined by a so-called scene graph, which is a data structure that arranges logical objects in form of a graph. Such objects are called *game objects* in Unity3D and refer to a set of components, like attached scripts, renderers or colliders. During runtime, the complete scene graph is rendered hierarchically once per frame while in each frame the properties of all game objects and their attached components are updated according to the current state of the scene graph. These modifications of game objects can be controlled by scripts, which can be written in the languages C#, JavaScript or Boo.

<sup>2</sup><http://unity3d.com/public-relations> (last accessed 27 August 2014)



The *Unity Asset Store* is a repository of existing components that are bundled in so-called *prefabs* and distributed in *unity packages*, which can be imported into any project. Developers can add their own projects or prefabs to the repository and, thus, make them publicly available for others, who can download and use those assets in their projects. When using assets, the time of development for a project can be reduced, especially when 3D models are needed. *KinectWalker* integrates several assets, which will also be described in this section.

Another important feature of Unity3D is the possibility of animating rigged character models by importing pre recorded animations from motion capturing (mocap) recordings. Moreover, the fact that animations can be influenced by modifying the underlying skeleton model of a character using scripts is important for the development of *KinectWalker*. In *KinectWalker*, a free character model from the Unity3D IDE was used as a humanoid wizard to help users calibrate the Kinect sensor and interact with the application. For that, motion capturing data was recorded and mapped onto the model. In addition, the character mimics the user's movements during runtime to provide visual feedback of recognized gestures. Thus, the user is able to see in realtime if his intended movements were captured correctly by the system.

Finally, Unity3D provides the possibility of extending the IDE via own plugins, which are defined by scripts. This feature was used to create the *LogDog* extension in order to enable visualizations of recorded data directly within the graphical user interface (GUI) of the Unity3D IDE.

### **Kinect with MS-SDK v1.10**

*Kinect with MS-SDK* is an asset by RF solutions, which contains a bundle of wrapper classes for the original Microsoft *Kinect for Windows SDK*. Wrapper classes define data structures that fulfill the role of a *facade* from the facade design pattern (Gamma, Helm, Johnson, & Vlissides, 1994) by wrapping the functionality of an underlying implementation and providing a reduced interface of it to requesting clients. The advantage of those classes is that they provide a new interface for the wrapped implementation which is independent from the original sources. In case of *Kinect with MS-SDK*, the asset wraps the Microsoft *Kinect for Windows SDK* and makes its functionality available within the Unity3D IDE. This allows *KinectWalker* to communicate with the native Microsoft SDK via the supported programming languages of Unity3D. It is mostly used to get the positional and rotational information from the 20 body joints that are tracked by the Kinect sensor. The *Kinect with MS-SDK* asset also offers a simple way of mapping this information onto game objects for each of the joints within the Unity3D IDE.

### Lounge Pack v1.0

For the training exercise, the *Lounge Pack* asset<sup>3</sup> was used as basis for the level design. Since users should learn to use the implemented interaction models within a realistic environment like a house, corresponding 3D models had to be created. By using this asset, the time of developing the level design for the training exercise could drastically be reduced, since the asset provides the necessary 3D models, textures and game objects. Only small modifications were needed in order to create the final level design.

### Cartoon FX Pack 3 v1.11

In order to make the visual landmarks within the exercises of *KinectWalker* more attractive for the users, the asset *Cartoon FX Pack 3*<sup>4</sup> by Jean Moreno was used. With this asset, the green light circles, which indicate targets for the navigation task in the *KinectWalker* application, were generated and animated.

### Cinema Mo Cap v1.1.1

The asset *Cinema Mo Cap*<sup>5</sup> by Cinema Suite Inc was used to create animations for the humanoid wizard of *KinectWalker*. By using the Kinect sensor, this asset creates motion capturing recordings of a person, which can later be attached to rigged body models in Unity3D. Thus, for the purpose of *KinectWalker* the PSI calibration gesture as well as the waving gesture for the interaction with the application were recorded and attached to the wizard.

## XAMPP

To persist relevant information during a training session of the presented user study from chapter 3, a MySQL database was used in combination with a local web server, providing a simple PHP interface to the *KinectWalker*. As framework, a common XAMPP<sup>6</sup> installation was set up. A detailed description of the webservice is given in section 4.3.3.

### 4.3.3. LogDog

*LogDog* has been developed as a stand-alone application, which is able to persist information from a Unity3D application on a MySQL database. The application is divided into two parts. The first part consists of several C# Unity3D scripts which can be used by Unity3D projects to communicate with an apache web server. Thus, *LogDog* can be

---

<sup>3</sup><https://www.assetstore.unity3d.com/en/#!/content/17085> (last accessed 20 August 2014)

<sup>4</sup><https://www.assetstore.unity3d.com/en/#!/content/10172> (last accessed 20 August 2014)

<sup>5</sup><https://www.assetstore.unity3d.com/en/#!/content/12620> (last accessed 20 August 2014)

<sup>6</sup><https://www.apachefriends.org/de/index.html> (last accessed 20 August 2014)

used to send information to the web server or to request information from the server in order to visualize it within the Unity3D IDE. For the latter feature, *LogDog* provides additional scripts, which extend the default Unity3D *EditorWindow* class in order to provide a GUI for the user (see fig. 4.5).

Within the *LogDog* GUI, a user can first of all import all recorded data from a database. After that, the user is able to select a participant from the according list and configure the way in which the walk path of this participant should be displayed. He can configure the width and colors for different objects, like simple waypoints, waypoints at which errors occurred, or those waypoints at which landmarks were selected. He can also display the viewing direction of the virtual avatar for each waypoint. The resulting walk path is then visualized within the scene editor of the Unity3D IDE, where the user can investigate it in detail. If he selects e.g. a waypoint at which an error occurred, like it is shown in fig. 4.5, he can see from the *object inspector*, which game object caused the error.

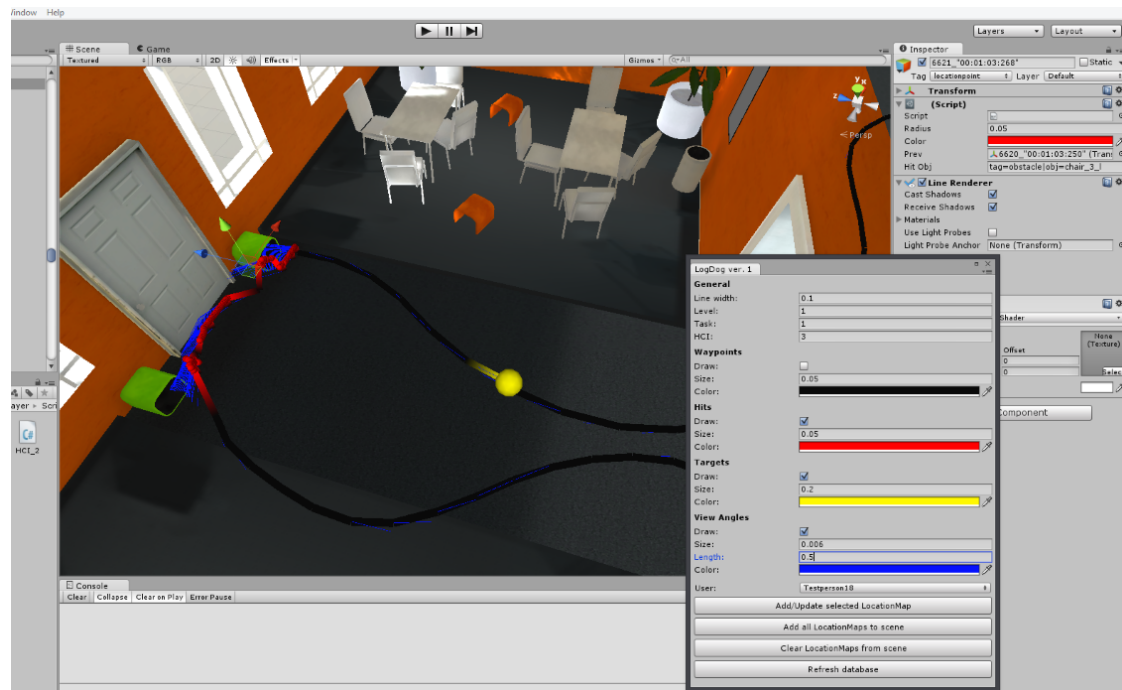


Figure 4.5.: The graphical user interface of *LogDog* in the Unity3D IDE. The image shows one part of the walk path of participant 18 in the training level. At the top right, one can see that the *game object* “chair\_3\_1” has caused the selected error.

The second component of *LogDog* is a web service, which runs on an apache web server. The web service consists of one class that handles the database connection and one class that handles incoming requests from a Unity3D application. Figure 4.6 shows the database model that describes how information is persisted. The web service is able to execute four different commands, each of which is defined by the GET parameter *action* of an incoming HTTP request.

- **log\_location:** This action instructs the server to store the given POST parameters into the table *locations*. This data consists of the positional and rotational information from the virtual avatar. In addition, information about the executed task is given, which covers the level number, the active interaction model and, if multiple tasks are defined, the number of the active task, as well as the ingame timestamp at which the HTTP request was initiated. Finally, also information about objects that were hit by the virtual avatar is logged. It is later used to decide whether the object is an obstacle which caused an error or a landmark which has been collected. Within the GUI, this information is visualized differently according to the color configuration that the user chooses.
- **add\_user:** This action is called at the beginning of a training session and instructs the server to store user-specific data in the table *users*. This data consists of a unique name for the user, his age or age class and his gender.
- **log\_performance:** When a level is completed, this action is sent to the server, including the performance data *time of error*, *number of errors* and *completion time*. As for locations, information about the level, the selected model and the active task is given in addition.
- **get\_all:** This action instructs the server to bundle all stored information into one JSON object and to send this object back to the requesting client.

The advantage of using a client-server architecture is that data can be stored on a centralized server. Regarding *KinectWalker*, this design allows the execution of training tasks at different locations, while all information is stored on the same database.

#### 4.3.4. Interaction Models

In the following, the implementation of the three interaction models LEAN, WiP and D2V of the *KinectWalker* application will be described. As shown in fig. 4.1, each of the models was implemented as a plugin to the *HCI Controller* component of *KinectWalker*. For that, every model must inherit the abstract class *HCI*, which expects the abstract functions *rotate*, *move* and *reset* to be implemented by its subclasses. While the function *reset* is mainly used to reset local initiated variables, the methods *move* and

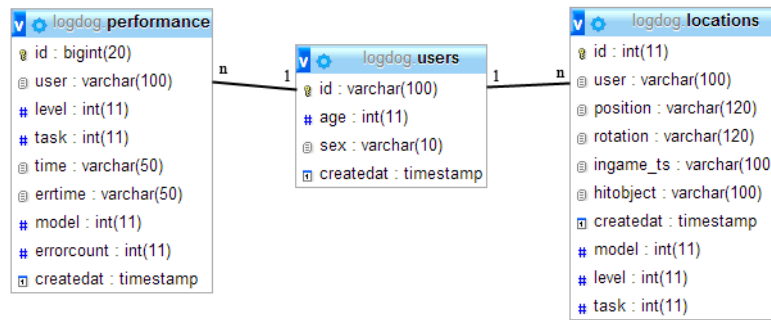


Figure 4.6.: UML diagram of the LogDog database model

*rotate* both get a reference to an instance of the class *AvatarController* from the *Kinect with MS-SDK* toolkit as parameter and directly modify the virtual avatar by moving and rotating it. The *AvatarController* is used as a provider to get the current positional and rotational information of all body joints which are being tracked by the Kinect sensor. This information covers a triple of floating point numbers to indicate the X-, Y- and Z-position of a joint within a three dimensional grid and a quaternion for the present rotation of that joint. The positional and rotational data is used by the different HCI subclasses to recognize gestures and to modify the virtual avatar accordingly. The complete process of how the virtual avatar is updated per rendering iteration of the scene graph is given in fig. 4.7.

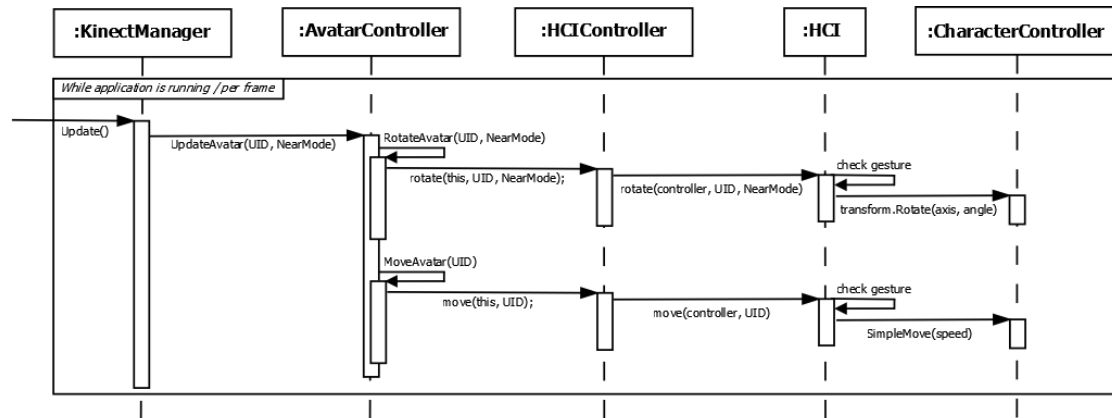


Figure 4.7.: Sequence diagram of the update process of the virtual avatar.

For the purpose of the user study in this thesis, only translational locomotion is important. For that, rotation was implemented in the same way for all interaction models. Furthermore, an additional configuration script was used to define offsets and limits for each interaction model.

## Rotation

As rotation was not the main focus of the evaluation of this thesis, it was modeled in the same way for all three interaction models. Therefore, a simple but intuitive interaction was used for rotating the virtual avatar. The center between the left and the right shoulder of a user is provided as an additional joint by the *Kinect with MS-SDK* asset. It represents the neck of a user. The rotation of the virtual avatar is computed by the rotation of the neck. This computation is defined by the difference of the angle around the vertical y-axis between the initial rotation of the neck and the one that is currently being tracked (see fig. 4.8). The absolute value of that difference then indicates the speed of the rotation, while the direction is based on the sign of the value.

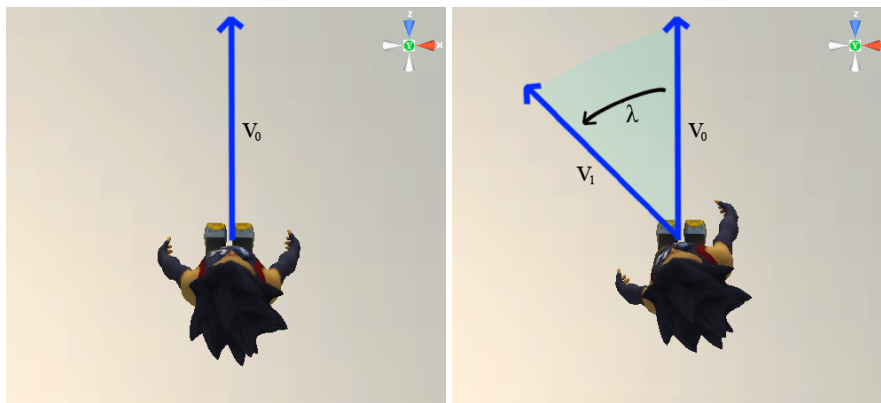


Figure 4.8.: Rotation gesture. On the left: the initial position of a user with vector  $v_0$  describing the initial direction orthogonal to both shoulders. On the right: the user after rotation. The rotation around the y-axis is indicated by  $\alpha$ .

## Model 1 - Lean

The Lean interaction model is based on the angle between the vertical y-axis and a line that goes through the abstract joint *neck* and the joint *waist* of a user (see fig. 4.9). Since this angle can also be described by the rotation around the x-axis of the Unity3D game object that is mapped to the joint *spine*, the computation only takes the latter into consideration. Thus, the movement is defined by the difference between the initial and the current rotation around the x-axis of the joint *spine*. The result is used to define the speed of a translational movement. Additionally, the model checks if the user's head is nearer to the Kinect sensor than his spine. In this case, the direction of the translational movement is assumed to be forward. If the spine is nearer to the sensor than the head, the direction is assumed to be backward.

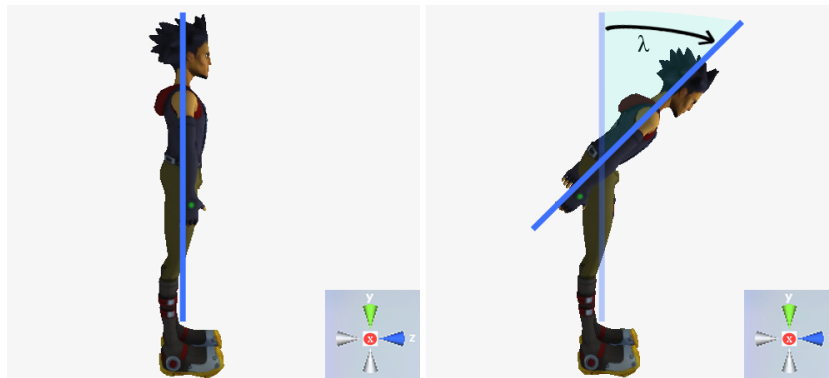


Figure 4.9.: Lean gesture. On the left: the initial position of a user. On the right: the user after having performed the LEAN gesture to go forward. The rotation around the x-axis is indicated by  $\alpha$ .

### Model 2 - Walking-In-Place

Walking-In-Place is, in contrast to LEAN and D2V, based on time-based interactions with the system, i.e. the user can influence the speed of translational movements by increasing the speed of repeating the walking gesture. Thus, the gesture is defined by the recognition of lifting the right or the left knee and further by a combination of newly recognized gestures with the last computed value.

As mentioned above, one part of the model is to recognize when a user lifted one of his knees. For that, a knee's position on the y-axis is compared to a configurable offset height. If the measured height is higher than the offset value, the system recognizes the lifting gesture (see fig. 4.10). An existing speed variable is then increased by a constant value. On each frame, the speed value is also decreased by being multiplied with the factor 0.986, which was defined by testing. As a result, the user can speed up the translational movement up to a configurable limit as he increases the speed of lifting his left and right knee alternately. The movement, however, slows down if he stops lifting his knees and it is stopped completely after a short time of inactivity.

### Model 3 - Distance-To-Velocity

For D2V, the movement speed is defined by the distance of the user's waist to its initial position, measured on the z- and x-axis. Thus, in the VR setup that was used in this thesis, the speed corresponds to the distance of a user to the center of the interaction field. As shown in fig. 4.11, the vector  $v_0$ , which describes the distance from the initial to the current position of the user's waist, is used to indicate the speed and the direction of the translational movement. Changes on the vertical v-axis were zeroed.

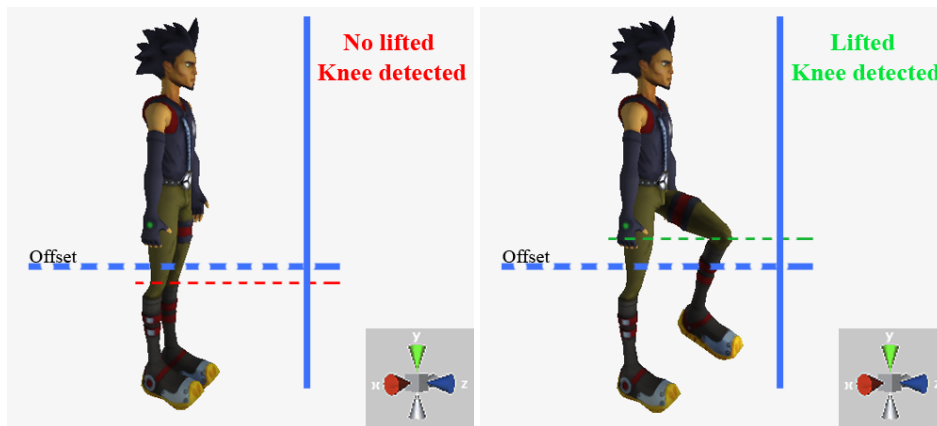


Figure 4.10.: Walking-In-Place gesture. On the left: no lifted knee is detected because none of the knees is higher than the offset. On the right: the left knee is lifted.



Figure 4.11.: Distance-To-Velocity gesture. On the left: the initial position of the user. On the right: the user has moved and the vector  $v_0$  describes the direction and speed.



## 4.4. Integration into SafeChild

One goal of this thesis was to integrate the interaction model which produces the lowest cognitive workload as a prototype into the SafeChild project. However, the results of the user study from chapter 3 show that none of the selected interaction models achieve good ratings for all kinds of users. It rather depends on *age*, *gender* and *VR experience*, which model performs most effectively in the context of a VR training application. Thus, the complete set of presented models was integrated into the SafeChild Unity3D project, including the possibility of switching between them during runtime.

The whole integration process was done in three steps. The first step already started during the development of *KinectWalker* as features like adapting the height of the virtual camera according to the user's body height or the support of a triple screen setup already focused on the SafeChild project. The second step was packaging and distributing *KinectWalker* to the developer of the SafeChild project, which was done by providing the source files as a bundled Unity package. Finally, in the third step, the developer was supported in adapting the source files of *KinectWalker* to the structure of the SafeChild project.

While the first two steps were done without problems, the third step revealed some issues in the implementations of the interaction models. Since the SafeChild setup uses the Microsoft Kinect for Windows sensor, some models were not responding with the same accuracy than in the setup of this thesis. This indicates that the hardware versions of the Kinect sensor might also affect the user-experience of locomotion interaction based on *KinectWalker*. Further versions of *KinectWalker* should, therefore, address these problems and should provide additional configurations for some of the integrated models.

## 5. Conclusion

This chapter concludes the thesis and summarizes the findings of the presented work. It will first give a short summary of what has been done in this thesis. After that, an outlook will be given, in which further developments and open research questions will be discussed.

### 5.1. Summary

As described in section 1.2, the goals of this thesis are to evaluate different natural user interfaces in the context of locomotion behavior in Virtual Reality applications. In detail, the correlation between their level of abstraction regarding real walking behavior and the cognitive workload which they produce should be investigated. One of the tested interaction models should then be implemented in the SafeChild project.

First, the terms *Virtual Reality* and *natural user interface* were explained. Then, a short overview of the *Cognitive Load Theory* of Chandler and Sweller (1991) was given. After this, a discussion about interaction models for the purpose of implementing natural locomotion interaction in a VR application was held, in which the three promising models *Lean* (LEAN), *Walking-In-Place* (WiP) and *Distance-To-Velocity* (D2V) were chosen to be evaluated in a user study.

Chapter 3 described the user study, in which the subjective perceived cognitive workload of the three models was tested with the NASA Task Load Index. These results were compared to objective measured performance data from a dual task exercise. With reference to the discussion about the models, the findings of the study were interpreted and concluded.

After the description of the study, the implementation of the *KinectWalker* application, which was developed for the purpose of this thesis, was explained. Its design and a detailed description of the gesture definitions for the three models LEAN, WiP and D2V were given in this chapter. It closed with an explanation about the integration of several models into the SafeChild project.

## 5.2. Findings

Referring to the discussion about existing interaction models, the three models LEAN, WiP and D2V can be sorted in the following order regarding their level of abstraction: LEAN is assumed to be most abstract in terms of natural walking behavior. WiP is assumed to be more natural than LEAN because it mimics real walking gestures in place. Nevertheless, it lacks spatial movements. D2V is assumed to be the most natural model among the selected three models.

The conducted user study, in which users had to perform a dual task exercise consisting of a memorizing task and a navigation task, revealed the following results:

According to the subjective data from the NASA TLX questionnaire, D2V was found to produce the lowest amount of cognitive load. LEAN performed slightly worse than D2V and WiP produced the highest perceived cognitive workload. This ranking does not match the assumed order regarding the level of abstraction. However, the answers to the questions about the *best* and the *worst* model of the post test questionnaire confirm the above-mentioned ranking.

It can also be concluded that *age* and *VR experience* might affect the perceived workload. Despite the fact that none of the results for the NASA TLX and for each interaction model differed significantly between the investigated groups, it could be observed that female participants perceived a significant higher *mental demand* than males when using the D2V model. Moreover, D2V was found to produce almost the same amount of cognitive workload as LEAN in the group of female participants. Nevertheless, for all evaluated groups of users, the above-mentioned ranking is valid, except for one group. VR inexperienced users preferred LEAN over D2V and WiP. This group even achieved the best results in the memorizing task when using the LEAN model compared to more VR experienced users. Thus, it can be concluded that VR inexperienced users perform best when using an abstract, unnatural interaction model.

The recorded walk path of the participants further show that most of the errors were made during the last part of the navigation task, after 6 of 10 navigation goals were reached. This shows that a high cognitive load also affects the navigation performance of users.

The fact that all findings from the conducted user study are based on a sample of only 22 participants must be taken into account, especially for the presented significance tests. Thus, to confirm the presented findings, further studies with a larger sample size have to be conducted.

## 5.3. Outlook

This section gives an overview of future work and developments that could be based on the findings of this thesis.

### 5.3.1. Evaluation

As mentioned before, the sample size of the conducted user study was too small to confirm all assumptions that were made. Therefore, user studies with more participants must be conducted. Besides the sample size, the distribution of demographic parameters could be considered more precisely. That means that a future study might investigate only the effect that the parameter *age* might have on the perceived workload by selecting a large and evenly distributed sample of people from several pre-defined age classes. Moreover, a study could investigate the effect of different levels of VR experience with the help of a pilot study being conducted in advance to find according samples.

VR training applications which are based on natural locomotion interaction using the Kinect sensor might, however, be developed according to the presented findings. Developers could e.g. integrate one of the presented interaction models depending on their target user group.

### 5.3.2. Implementation

Developed for the purpose of being used as a toolkit within the user study of this thesis, *KinectWalker* provides an open and extendable design for integrating natural interaction models into any Unity3D application based on the Kinect sensor. As described in this work, the application can be used to conduct user studies or to integrate natural interaction models into a real VR application easily. The latter feature allows developers to reduce their time of development when implementing natural locomotion interaction.

In the future development of the *KinectWalker* application, more pre-defined interaction models could be implemented and provided to other VR developers, and the presented three models might be improved. It would also be possible to distribute the application as a Unity3D asset within the Unity asset store.

The *LogDog* application was used to visualize performance data of the presented user study within the Unity3D IDE. If distributed via the asset store, the *LogDog* editor extension might be used by other researchers to help them visualize recorded data of VR navigation tasks. This might, in return, result in new findings for the research in VR.

# A. Appendix

## A.1. Pre Test Questionnaire

Testperson: \_\_\_\_

Wie alt sind Sie?

10 bis 14     15 bis 20     21 bis 40     Älter als 40

Haben Sie in den letzten 4 Jahren ein Computerspiel gespielt?

Ja     Nein

Um welche Art von Computerspiel handelte es sich dabei?

(Nur beantworten, wenn die vorherige Frage mit „Ja“ beantwortet wurde)

*Mehrfachauswahl möglich!*

Egoshooter     Adventure     Strategie     Simulation  
 Sonstige: \_\_\_\_\_

Besitzen Sie einen Kinect Sensor?

Ja     Nein

Nutzen Sie einen Computer zum Lernen?

Ja     Nein

Wie nutzen Sie den Computer zum Lernen?

(Nur beantworten, wenn die vorherige Frage mit "Ja" beantwortet wurde)

*Mehrfachauswahl möglich*

Internetrecherche     Serious Games / Lernspiele  
 eLearning Software (z.B. Vokabeltrainer)  
 Hilfsprogramme (z.B. Excel, Matlab, usw.)  
 Sonstige: \_\_\_\_\_

Wie lange nutzen Sie ihren Computer durchschnittlich am Tag?

0 Stunden     1 Stunde     2 Stunden     3 Stunden  
 Länger als 3 Stunden     Ich besitze keine Computer

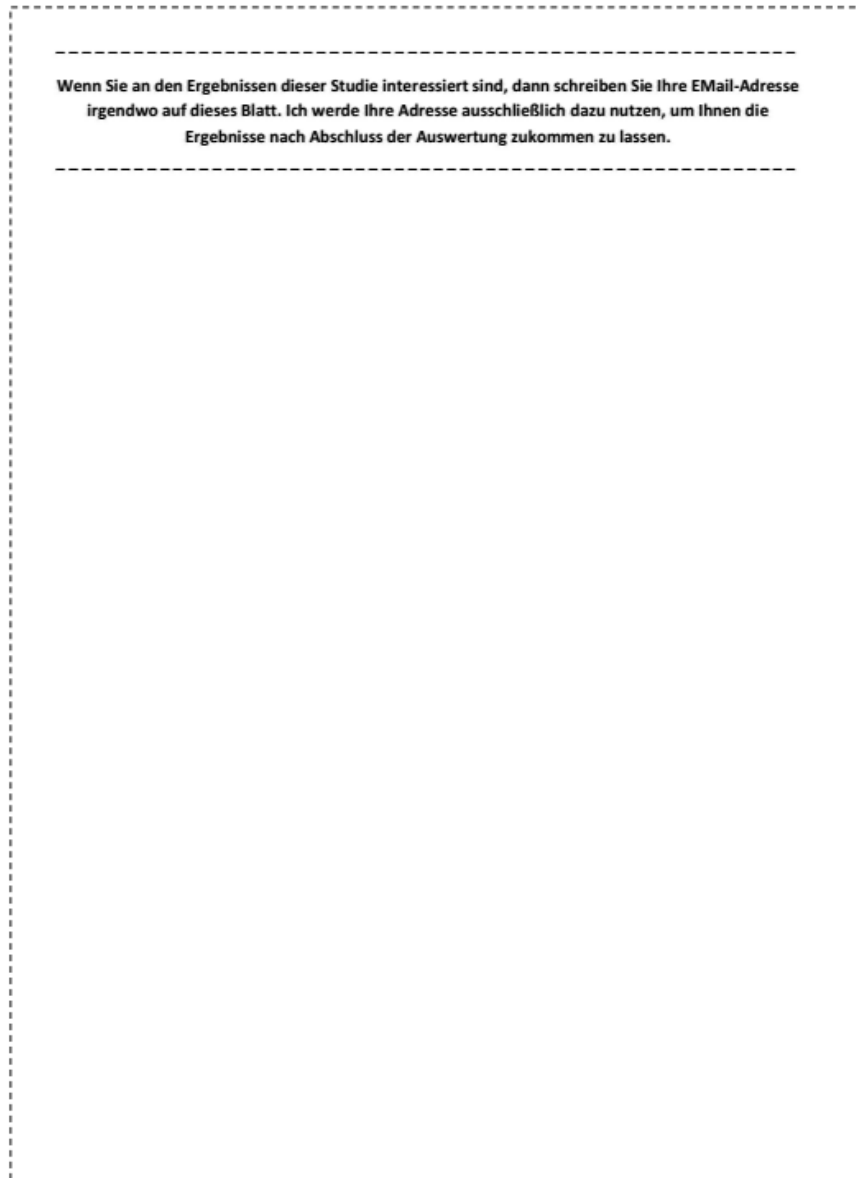
## A.2. Email Response Sheet

The sheet that was offered to the participants of the user study on which they could write down their Email address in order to be informed about the results of the thesis.

-----

**Wenn Sie an den Ergebnissen dieser Studie interessiert sind, dann schreiben Sie Ihre EMail-Adresse irgendwo auf dieses Blatt. Ich werde Ihre Adresse ausschließlich dazu nutzen, um Ihnen die Ergebnisse nach Abschluss der Auswertung zukommen zu lassen.**

-----





## A.3.2. Second part of NASA TLX

Testperson: \_\_\_\_ HCI: \_\_\_\_

Lösungstext: \_\_\_\_\_

Entscheiden Sie sich bitte bei jeder der 15 nachfolgenden Fragen für die Dimension, die für das Gesamtempfinden hinsichtlich der Übung für Sie die jeweils bedeutsamer war.

Geistige Anforderung	<input type="checkbox"/>	<input type="checkbox"/>	Frustration
Geistige Anforderung	<input type="checkbox"/>	<input type="checkbox"/>	Zeitliche Anforderung
Leistung	<input type="checkbox"/>	<input type="checkbox"/>	Frustration
Frustration	<input type="checkbox"/>	<input type="checkbox"/>	Anstrengung
Geistige Anforderung	<input type="checkbox"/>	<input type="checkbox"/>	Körperliche Anforderung
Anstrengung	<input type="checkbox"/>	<input type="checkbox"/>	Körperliche Anforderung
Anstrengung	<input type="checkbox"/>	<input type="checkbox"/>	Leistung
Frustration	<input type="checkbox"/>	<input type="checkbox"/>	Körperliche Anforderung
Körperliche Anforderung	<input type="checkbox"/>	<input type="checkbox"/>	Leistung
Leistung	<input type="checkbox"/>	<input type="checkbox"/>	Zeitliche Anforderung
Geistige Anforderung	<input type="checkbox"/>	<input type="checkbox"/>	Anstrengung
Zeitliche Anforderung	<input type="checkbox"/>	<input type="checkbox"/>	Körperliche Anforderung
Leistung	<input type="checkbox"/>	<input type="checkbox"/>	Geistige Anforderung
Zeitliche Anforderung	<input type="checkbox"/>	<input type="checkbox"/>	Frustration
Zeitliche Anforderung	<input type="checkbox"/>	<input type="checkbox"/>	Anstrengung



### A.3.3. Information Sheet

Optional sheet that was handed to the users when filling out the second part of NASA TLX questionnaire.

#### Geistige Anforderung

Wieviel geistige Anforderung war bei der Informationsaufnahme (Hören der Zahlen, Suchen des nächsten Lichtkreises) und bei der Informationsverarbeitung (Merken der Zahlen, Bewegung zum nächsten Lichtkreis) erforderlich? War die Übung leicht oder anspruchsvoll, einfach oder komplex?

#### Zeitliche Anforderung

Wieviel Zeitdruck empfanden Sie? War die Übung langsam und geruhsam oder schnell und hektisch?

#### Leistung

Wie erfolgreich haben Sie Ihrer Meinung nach die vom Versuchsleiter (oder Ihnen selbst) gesetzten Ziele erreicht? Wie zufrieden waren Sie mit Ihrer Leistung?

#### Frustration

Wie unsicher, entmutigt, irritiert, gestresst und verärgert (bzw. sicher, bestätigt, zufrieden, entspannt und zufrieden mit sich selbst) fühlten Sie sich während der Übung?

#### Körperliche Anforderung

Wieviel körperliche Aktivität war erforderlich? War die Übung leicht oder schwer, einfach oder anstrengend, erholsam oder mühselig?

#### Anstrengung

Wie hart mussten Sie arbeiten, um Ihren Grad an Aufgabenerfüllung zu erreichen (geistig und körperlich)?

## A.4. Post Test Questionnaire

Testperson: \_\_\_\_\_

**Welche Interaktion gefiel Ihnen persönlich am besten?**

1. Lehnen       2. Laufen       3. Kreis

**Welche Interaktion gefiel Ihnen persönlich am wenigsten?**

1. Lehnen       2. Laufen       3. Kreis

**Haben Sie Verbesserungsvorschläge oder gab es in Ihren Augen einen Aspekt der Interaktionsmodelle, den Sie sich in auch in einer anderen Kombination vorstellen könnten?**  
(optional)

## A.5. Results of the User Study

On the following pages, the following abbreviations are used:

<b>C</b>	Completion time (in seconds)
<b>E</b>	Time of error (in seconds)
<b>#E</b>	Number of errors that occurred during the main task
<b>E%</b>	Relation between time of error and completion time
<b>R</b>	Result of the first part of NASA TLX on a 20-point scale
<b>R-N</b>	The normalized rating of <i>R</i> on a 100-point scale using 5-point steps
<b>hits</b>	Number that indicates how often a dimension was chosen in the second part of NASA TLX
<b>%</b>	The weighted factor for a dimension of NASA TLX
<b>#Prefix</b>	Number of correct answers regarding the method <i>correct digits in prefix</i>
<b>#Position</b>	Number of correct answers regarding the method <i>correct digits at position</i>

Moreover, the NASA TLX values for each model are marked in red, if the value is the worst of the three tested models. They are marked in green, if the according value was the best for the three models. In any other case, the value is marked in yellow.

Test person		1		2		3	
HCl		"Learning"		"Walking-In-Place"		"Distance-To-Velocity"	
C	E	#E	%	C	E	#E	%
75.4	30.6	3	14.1	71.9	30.0	3	41.7
12	60	5	0.333	15	75	5	0.333
10	50	3	0.200	11	55	1	0.067
8	40	3	0.200	17	85	2	0.133
9	45	3	0.200	17	85	4	0.267
12	60	0	0.000	17	85	2	0.133
<b>Overall</b>			<b>48.6667</b>				<b>77.6667</b>

Test person		1		2		3	
HCl		"Learning"		"Walking-In-Place"		"Distance-To-Velocity"	
C	E	#E	%	C	E	#E	%
83.1	5.2	2	8.4	107.2	0.0	0	0.0
4	20	3	0.200	9	45	5	0.333
7	35	2	0.133	8	40	2	0.133
3	15	0	0.000	5	25	0	0.000
11	55	5	0.333	8	40	4	0.267
3	15	1	0.067	9	45	1	0.067
7	35	4	0.267	7	35	3	0.200
<b>Overall</b>			<b>37.3333</b>				<b>43.0000</b>

Test person		1		2		3	
HCl		"Learning"		"Walking-In-Place"		"Distance-To-Velocity"	
C	E	#E	%	C	E	#E	%
80.5	8.5	5	10.6	122.4	17.1	6	14.0
16	80	5	0.333	18	90	2	0.133
1	5	0	0.000	17	85	0	0.000
8	40	4	0.267	4	20	4	0.267
15	75	3	0.200	14	70	4	0.267
13	65	2	0.133	17	85	2	0.133
5	25	1	0.067	19	95	3	0.200
<b>Overall</b>			<b>62.6667</b>				<b>66.3333</b>

Test person		1		2		3	
HCl		"Learning"		"Walking-In-Place"		"Distance-To-Velocity"	
C	E	#E	%	C	E	#E	%
83.0	6.7	1	8.1	120.6	0.0	0	0.0
15	75	1	0.067	9	45	3	0.200
5	25	5	0.333	9	45	4	0.267
10	50	2	0.133	5	25	1	0.067
10	50	4	0.267	5	25	5	0.333
12	60	3	0.200	3	15	2	0.133
2	10	0	0.000	2	10	0	0.000
<b>Overall</b>			<b>45.3333</b>				<b>33.0000</b>

Demographic data [i.e. answers from the pre test]		Age [21-40]	
Played computer games?	no	Has got a Kinect Sensor?	no
Uses PC for learning?	yes	How? Internet / Helpbooks	2
PC usage per day (hours)	2	Answers from post-test	
Feedback	Drehen war unständiglich / Rückwärts fehlt bei der zweiten Steuerung / Bremsen im Kreis schwierig	Best HCl	1
Worst HCl	2		

Demographic data [i.e. answers from the pre test]		Age [21-40]	
Played computer games?	yes	Has got a Kinect Sensor?	no
Uses PC for learning?	yes	How? Internet / Helpbooks	3
PC usage per day (hours)	>3	Answers from post-test	
Feedback	Lehnen: Zu viel Bewegung benötigt (anstrengend)	Best HCl	3
Worst HCl	1		

Demographic data [i.e. answers from the pre test]		Age [10-14]	
Played computer games?	yes	Has got a Kinect Sensor?	yes
Uses PC for learning?	no	How?	
PC usage per day (hours)	2	Answers from post-test	
Feedback	Vorschlag: Kreis + Laufen	Best HCl	3
Worst HCl	2		

Demographic data [i.e. answers from the pre test]		Age [10-14]	
Played computer games?	yes	Has got a Kinect Sensor?	yes
Uses PC for learning?	yes	How? Learning	2,5
PC usage per day (hours)	2,5	Answers from post-test	
Feedback	Testszenario ist langweilig	Best HCl	2
Worst HCl	1		

Demographic data [i.e. answers from the pre test]		Age [21-40]	
Played computer games?	yes	Has got a Kinect Sensor?	no
Uses PC for learning?	yes	How? Internet / Helpbooks	3
PC usage per day (hours)	2	Answers from post-test	
Feedback	Laufen: Zu viel Beinhebung, rückwärts nicht möglich	Best HCl	3
Worst HCl	1		

Demographic data [i.e. answers from the pre test]		Age [10-14]	
Played computer games?	yes	Has got a Kinect Sensor?	yes
Uses PC for learning?	no	How?	
PC usage per day (hours)	2	Answers from post-test	
Feedback	Vorschlag: Kreis + Laufen	Best HCl	3
Worst HCl	2		

Demographic data [i.e. answers from the pre test]		Age [10-14]	
Played computer games?	yes	Has got a Kinect Sensor?	yes
Uses PC for learning?	yes	How? Learning	4
PC usage per day (hours)	2,5	Answers from post-test	
Feedback	Testszenario ist langweilig	Best HCl	2
Worst HCl	1		

Demographic data [i.e. answers from the pre test]		Age [21-40]	
Played computer games?	no	Has got a Kinect Sensor?	no
Uses PC for learning?	yes	How? Internet / Helpbooks	2
PC usage per day (hours)	2	Answers from post-test	
Feedback	Drehen war unständiglich / Rückwärts fehlt bei der zweiten Steuerung / Bremsen im Kreis schwierig	Best HCl	1
Worst HCl	2		

Demographic data [i.e. answers from the pre test]		Age [10-14]	
Played computer games?	yes	Has got a Kinect Sensor?	yes
Uses PC for learning?	yes	How? Learning	4
PC usage per day (hours)	2,5	Answers from post-test	
Feedback	Testszenario ist langweilig	Best HCl	2
Worst HCl	1		

Demographic data [i.e. answers from the pre test]		Age [21-40]	
Played computer games?	no	Has got a Kinect Sensor?	no
Uses PC for learning?	yes	How? Internet / Helpbooks	2
PC usage per day (hours)	2	Answers from post-test	
Feedback	Drehen war unständiglich / Rückwärts fehlt bei der zweiten Steuerung / Bremsen im Kreis schwierig	Best HCl	1
Worst HCl	2		

Test person	HCl	1					2					3					
		"Learning"					"Walking-In-Place"					"Distance-To-Velocity"					
Measured Performance		C	E	#E	%	C	E	#E	%	C	E	#E	%	C	E	#E	%
NASA TLX		112.5	43.7	7	38.8	77.6	7.7	1	9.9	58.0	0.0	0	0.0	0	0.0		
Mental Demand		15	75	5	0.333	13	65	5	0.333	5	25	2	0.133	5	25	2	0.133
Physical Demand		3	15	2	0.133	2	10	1	0.067	2	10	5	0.333	1	5	0	0.000
Temporal Demand		1	5	0	0.000	1	5	0	0.000	1	5	0	0.000	1	5	0	0.000
Performance		7	35	3	0.200	5	25	4	0.267	4	20	4	0.267	4	20	4	0.267
Effort		3	15	2	0.133	5	25	2	0.133	1	5	2	0.133	1	5	2	0.133
Frustration		15	75	3	0.200	1	5	3	0.200	1	5	2	0.133	1	5	2	0.133
Overall		<b>51.0000</b>					<b>33.3333</b>					<b>13.3333</b>					

Test person	HCl	6					2					3					
		"Learning"					"Walking-In-Place"					"Distance-To-Velocity"					
Measured Performance		C	E	#E	%	C	E	#E	%	C	E	#E	%	C	E	#E	%
NASA TLX		97.4	11.9	5	12.2	86.6	0.0	0	0.0	67.7	0.1	1	0.1	0	0.0		
Mental Demand		3	15	3	0.200	1	5	2	0.133	1	5	4	0.267	1	5	4	0.267
Physical Demand		1	5	3	0.200	10	50	4	0.267	10	50	2	0.133	2	10	2	0.133
Temporal Demand		1	5	1	0.067	1	5	1	0.067	3	15	1	0.067	3	15	1	0.067
Performance		5	25	4	0.267	2	10	4	0.267	1	5	5	0.333	5	25	3	0.200
Effort		5	25	4	0.267	3	15	4	0.267	7	35	3	0.200	7	35	3	0.200
Frustration		1	5	0	0.000	3	15	0	0.000	1	5	0	0.000	1	5	0	0.000
Overall		<b>17.6667</b>					<b>21.0000</b>					<b>17.6667</b>					

Test person	HCl	7					2					3					
		"Learning"					"Walking-In-Place"					"Distance-To-Velocity"					
Measured Performance		C	E	#E	%	C	E	#E	%	C	E	#E	%	C	E	#E	%
NASA TLX		124.8	27.9	6	16.0	172.5	3.4	2	2.0	89.5	7.8	2	8.7	2	8.7		
Mental Demand		6	30	3	0.200	20	100	2	0.133	1	5	3	0.200	1	5	3	0.200
Physical Demand		13	65	1	0.067	20	100	4	0.267	1	5	3	0.200	1	5	3	0.200
Temporal Demand		10	50	0	0.000	1	5	0	0.000	1	5	0	0.000	1	5	0	0.000
Performance		10	50	3	0.200	20	100	2	0.133	1	5	4	0.267	1	5	4	0.267
Effort		10	50	4	0.267	20	100	2	0.133	1	5	4	0.267	1	5	4	0.267
Frustration		18	90	4	0.267	20	100	5	0.333	1	5	1	0.067	1	5	1	0.067
Overall		<b>57.6667</b>					<b>100.0000</b>					<b>5.0000</b>					

Test person	HCl	8					2					3					
		"Learning"					"Walking-In-Place"					"Distance-To-Velocity"					
Measured Performance		C	E	#E	%	C	E	#E	%	C	E	#E	%	C	E	#E	%
NASA TLX		98.0	1.0	1	1.0	78.0	3.5	1	4.5	48.8	0.0	0	0.0	0	0.0		
Mental Demand		3	15	4	0.267	4	20	2	0.133	17	85	4	0.267	4	20	4	0.267
Physical Demand		5	25	0	0.000	13	65	5	0.333	14	70	4	0.267	7	35	4	0.267
Temporal Demand		2	10	1	0.067	12	60	0	0.000	12	60	0	0.000	12	60	0	0.000
Performance		5	25	3	0.200	8	40	3	0.200	4	20	2	0.133	4	20	2	0.133
Effort		5	25	3	0.200	14	70	4	0.267	14	70	4	0.267	14	70	4	0.267
Frustration		2	10	3	0.200	5	25	1	0.067	7	35	1	0.067	7	35	1	0.067
Overall		<b>17.0000</b>					<b>52.6667</b>					<b>65.0000</b>					

Demographic data [i.e. answers from the pre test]		Age 10-14		Memory-Task	
Played computer games?	Yes			HCl	1
Type of games	Ego / Adv. / Sim.			Original	2
Has got a Kinect Sensor?	Yes			User	280724852
Uses PC for learning?	no			#Prefik	49872154_
How?				#Position	3
PC usage per day (hours)	1				4
Answers from post-test					2
Feedback					
Best HCl	3				
Worst HCl	1				

Demographic data [i.e. answers from the pre test]		Age 10-14		Memory-Task	
Played computer games?	Yes			HCl	1
Type of games	Ego / Adv. / Strat. / Sim.			Original	2
Has got a Kinect Sensor?	Yes			User	201352398
Uses PC for learning?	Yes			#Prefik	970125604
How?	relearning			#Position	5
PC usage per day (hours)	2				7
Answers from post-test					7
Feedback					
Best HCl	3				
Worst HCl	1				

Demographic data [i.e. answers from the pre test]		Age 21-40		Memory-Task	
Played computer games?	Yes			HCl	1
Type of games	Ego / Adv. / Strat.			Original	2
Has got a Kinect Sensor?	no			User	926590831
Uses PC for learning?	Yes			#Prefik	7
How?	Internet / elearning / Helpcools			#Position	8
PC usage per day (hours)	3				7
Answers from post-test					9
Feedback					
Best HCl	3				
Worst HCl	2				

Demographic data [i.e. answers from the pre test]		Age 21-40		Memory-Task	
Played computer games?	Yes			HCl	1
Type of games	Ego / Strat. / Sim.			Original	2
Has got a Kinect Sensor?	no			User	468014376
Uses PC for learning?	Yes			#Prefik	5
How?	Internet / Helpcools			#Position	5
PC usage per day (hours)	3				8
Answers from post-test					9
Feedback					
Best HCl	1				
Worst HCl	2				

Test person	HCl	5					2					3					
		"Learning"					"Walking-In-Place"					"Distance-To-Velocity"					
Measured Performance		C	E	#E	%	C	E	#E	%	C	E	#E	%	C	E	#E	%
NASA TLX		112.5	43.7	7	38.8	77.6	7.7	1	9.9	58.0	0.0	0	0.0	0	0.0		
Mental Demand		15	75	5	0.333	13	65	5	0.333	5	25	2	0.133	5	25	2	0.133
Physical Demand		3	15	2	0.133	2	10	1	0.067	2	10	5	0.333	1	5	0	0.000
Temporal Demand		1	5	0	0.000	1	5	0	0.000	1	5	0	0.000	1	5	0	0.000
Performance		7	35	3	0.200	5	25	4	0.267	4	20	4	0.267	4	20	4	0.267
Effort		3	15	2	0.133	5	25	2	0.133	1	5	2	0.133	1	5	2	0.133
Frustration		15	75	3	0.200	1	5	3	0.200	1	5	2	0.133	1	5	2	0.133
Overall		<b>51.0000</b>					<b>33.3333</b>					<b>13.3333</b>					

Test person	HCl	6					2					3					
		"Learning"					"Walking-In-Place"					"Distance-To-Velocity"					
Measured Performance		C	E	#E	%	C	E	#E	%	C	E	#E	%	C	E	#E	%
NASA TLX		97.4	11.9	5	12.2	86.6	0.0	0	0.0	67.7	0.1	1	0.1	0	0.0		
Mental Demand		3	15	3	0.200	1	5	2	0.133	1	5	4	0.267	1	5	4	0.267
Physical Demand		1	5	3	0.200	10	50	4	0.267	10	50	2	0.133	2	10	2	0.133
Temporal Demand		1	5	1	0.067	1	5	1	0.067	3	15	1	0.067	3	15	1	0.067
Performance		5	25	4	0.267	2	10	4	0.267	1	5	5	0.333	5	25	3	0.200
Effort		5	25	4	0.267	3	15	4	0.267	7	35	3	0.200	7	35	3	0.200
Frustration		1	5	0	0.000	3	15	0	0.000	1	5	0	0.000	1	5	0	0.000
Overall		<b>17.6667</b>					<b>21.0000</b>					<b>17.6667</b>					

Test person	HCl	7					2					3					
		"Learning"					"Walking-In-Place"					"Distance-To-Velocity"					
Measured Performance		C	E	#E	%	C	E	#E	%	C	E	#E	%	C	E	#E	%
NASA TLX		124.8	27.9	6	16.0	172.5	3.4	2	2.0	89.5	7.8	2	8.7	2	8.7		
Mental Demand		6	30	3	0.200	20	100	2	0.133	1	5	3	0.200	1	5	3	0.200
Physical Demand		13	65	1	0.067	20	100	4	0.267	1	5	3	0.200	1	5	3	0.200
Temporal Demand		10	50	0	0.000	1	5	0	0.000	1	5	0	0.000	1	5	0	0.000
Performance		10	50	3	0.200	20	100	2	0.133	1	5	4	0.267	1	5	4	0.267
Effort		10	50	4	0.267	20	100	2	0.133	1	5	4	0.267	1	5	4	0.267
Frustration		18	90	4	0.267	20	100	5	0.333	1	5	1	0.067	1	5	1	0.067
Overall		<b>57.6667</b>					<b>100.0000</b>					<b>5.0000</b>					

Test person	HCl	8					2					3					
		"Learning"					"Walking-In-Place"					"Distance-To-Velocity"					
Measured Performance		C	E	#E	%	C	E	#E	%	C	E	#E	%	C	E	#E	%
NASA TLX		98.0	1.0	1	1.0	78.0	3.5	1	4.5	48.8	0.0	0	0.0	0	0.0		
Mental Demand		3	15	4	0.267	4	20	2	0.133	17	85	4	0.267	4	20	4	0.267
Physical Demand		5	25	0	0.000	13	65	5	0.333	14	70	4	0.267	7	35	4	0.267
Temporal Demand		2	10	1	0.067	12	60	0	0.000	12	60	0	0.000	12	60	0	0.000
Performance		5</															

Test person		9		2		3	
HCl		"Learning"		"Walking-In-Place"		"Distance-To-Velocity"	
C	E	#E	E%	C	E	#E	E%
102.2	11.3	5	11.1	72.0	0.0	0	0.0
168.3	24.9	3	15.7	108.4	0.0	0	0.0
17	85	4	0.267	15	75	2	0.133
2	10	4	0.267	9	45	5	0.333
5	25	1	0.067	5	25	1	0.067
16	80	3	0.200	12	60	3	0.200
8	40	0	0.000	10	50	0	0.000
12	60	5	0.333	14	70	2	0.133
7	35	2	0.133	13	65	5	0.333
10	50	1	0.067	11	55	3	0.200
<b>Overall</b>			<b>44.0000</b>				<b>37.6667</b>

Test person		10		3		9	
HCl		"Learning"		"Walking-In-Place"		"Distance-To-Velocity"	
C	E	#E	E%	C	E	#E	E%
102.2	11.3	5	11.1	72.0	0.0	0	0.0
168.3	24.9	3	15.7	108.4	0.0	0	0.0
17	85	4	0.267	15	75	2	0.133
2	10	4	0.267	9	45	5	0.333
5	25	1	0.067	5	25	1	0.067
16	80	3	0.200	12	60	3	0.200
8	40	0	0.000	10	50	0	0.000
12	60	5	0.333	14	70	2	0.133
7	35	2	0.133	13	65	5	0.333
10	50	1	0.067	11	55	3	0.200
<b>Overall</b>			<b>59.0000</b>				<b>58.6667</b>

Test person		11		2		3	
HCl		"Learning"		"Walking-In-Place"		"Distance-To-Velocity"	
C	E	#E	E%	C	E	#E	E%
102.2	11.3	5	11.1	72.0	0.0	0	0.0
168.3	24.9	3	15.7	108.4	0.0	0	0.0
17	85	4	0.267	15	75	2	0.133
2	10	4	0.267	9	45	5	0.333
5	25	1	0.067	5	25	1	0.067
16	80	3	0.200	12	60	3	0.200
8	40	0	0.000	10	50	0	0.000
12	60	5	0.333	14	70	2	0.133
7	35	2	0.133	13	65	5	0.333
10	50	1	0.067	11	55	3	0.200
<b>Overall</b>			<b>57.3333</b>				<b>48.0000</b>

Test person		12		2		3	
HCl		"Learning"		"Walking-In-Place"		"Distance-To-Velocity"	
C	E	#E	E%	C	E	#E	E%
102.2	11.3	5	11.1	72.0	0.0	0	0.0
168.3	24.9	3	15.7	108.4	0.0	0	0.0
17	85	4	0.267	15	75	2	0.133
2	10	4	0.267	9	45	5	0.333
5	25	1	0.067	5	25	1	0.067
16	80	3	0.200	12	60	3	0.200
8	40	0	0.000	10	50	0	0.000
12	60	5	0.333	14	70	2	0.133
7	35	2	0.133	13	65	5	0.333
10	50	1	0.067	11	55	3	0.200
<b>Overall</b>			<b>80.6667</b>				<b>75.0000</b>

Test person		1		2		3	
HCl		Age [21-40]		Memory-Task		Memory-Task	
Original	#Prefik	Original	#Prefik	Original	#Prefik	Original	#Prefik
594173082	791045	791045782	791045	791045782	791045	791045782	791045
6947708	791045	6947708	791045	6947708	791045	6947708	791045
#Position	3	#Position	3	#Position	3	#Position	3

Test person		1		2		3	
HCl		Age [21-40]		Memory-Task		Memory-Task	
Original	#Prefik	Original	#Prefik	Original	#Prefik	Original	#Prefik
149627031	59431028	149627031	59431028	149627031	59431028	149627031	59431028
14927031	59431028	14927031	59431028	14927031	59431028	14927031	59431028
#Position	3	#Position	3	#Position	3	#Position	3

Test person		1		2		3	
HCl		Age [21-40]		Memory-Task		Memory-Task	
Original	#Prefik	Original	#Prefik	Original	#Prefik	Original	#Prefik
17892546	706415728	17892546	706415728	17892546	706415728	17892546	706415728
178026	706415728	178026	706415728	178026	706415728	178026	706415728
#Position	4	#Position	4	#Position	4	#Position	4

Test person		1		2		3	
HCl		Age [21-40]		Memory-Task		Memory-Task	
Original	#Prefik	Original	#Prefik	Original	#Prefik	Original	#Prefik
149627031	59431028	149627031	59431028	149627031	59431028	149627031	59431028
14927031	59431028	14927031	59431028	14927031	59431028	14927031	59431028
#Position	3	#Position	3	#Position	3	#Position	3

Test person		1		2		3	
HCl		Age [21-40]		Memory-Task		Memory-Task	
Original	#Prefik	Original	#Prefik	Original	#Prefik	Original	#Prefik
17925640	217450861	17925640	217450861	17925640	217450861	17925640	217450861
170325640	217450861	170325640	217450861	170325640	217450861	170325640	217450861
#Position	2	#Position	2	#Position	2	#Position	2

Test person		1		2		3	
HCl		Age [21-40]		Memory-Task		Memory-Task	
Original	#Prefik	Original	#Prefik	Original	#Prefik	Original	#Prefik
17925640	217450861	17925640	217450861	17925640	217450861	17925640	217450861
170325640	217450861	170325640	217450861	170325640	217450861	170325640	217450861
#Position	2	#Position	2	#Position	2	#Position	2

Test person		1		2		3	
HCl		Age [21-40]		Memory-Task		Memory-Task	
Original	#Prefik	Original	#Prefik	Original	#Prefik	Original	#Prefik
17925640	217450861	17925640	217450861	17925640	217450861	17925640	217450861
170325640	217450861	170325640	217450861	170325640	217450861	170325640	217450861
#Position	2	#Position	2	#Position	2	#Position	2

Test person		1		2		3	
HCl		Age [21-40]		Memory-Task		Memory-Task	
Original	#Prefik	Original	#Prefik	Original	#Prefik	Original	#Prefik
17925640	217450861	17925640	217450861	17925640	217450861	17925640	217450861
170325640	217450861	170325640	217450861	170325640	217450861	170325640	217450861
#Position	2	#Position	2	#Position	2	#Position	2

Test person		1		2		3	
HCl		Age [21-40]		Memory-Task		Memory-Task	
Original	#Prefik	Original	#Prefik	Original	#Prefik	Original	#Prefik
17925640	217450861	17925640	217450861	17925640	217450861	17925640	217450861
170325640	217450861	170325640	217450861	170325640	217450861	170325640	217450861
#Position	2	#Position	2	#Position	2	#Position	2

Test person		1		2		3	
HCl		Age [21-40]		Memory-Task		Memory-Task	
Original	#Prefik	Original	#Prefik	Original	#Prefik	Original	#Prefik
17925640	217450861	17925640	217450861	17925640	217450861	17925640	217450861
170325640	217450861	170325640	217450861	170325640	217450861	170325640	217450861
#Position	2	#Position	2	#Position	2	#Position	2

Demographic data [i.e. answers from the pre test]												
Age [21-40]												
Played computer games?	Yes	1	2	3								
Type of games	Exp/Adv./Sim./Other (Pbils., Sport)	721509384	721509384	970381624								
Has got a Kinect Sensor?	no	483671059	721509384	90738124								
Uses PC for learning?	yes	5	5	1								
How?	Internet / Helpbooks	6	7	4								
PC usage per day (hours)	>3											
Answers from post-test												
Feedback												
Best HCI	3											
Worst HCI	1											

Demographic data [i.e. answers from the pre test]												
Age [21-40]												
Played computer games?	Yes	1	2	3								
Type of games	Sim.	627045189	549164901	549164901								
Has got a Kinect Sensor?	no	968251613	620745689	549164901								
Uses PC for learning?	yes	6	5	5								
How?	Internet											
PC usage per day (hours)	>3											
Answers from post-test												
Feedback												
Best HCI	3											
Worst HCI	1											

Demographic data [i.e. answers from the pre test]												
Age [10-14]												
Played computer games?	Yes	1	2	3								
Type of games	Adv. / Arcade	130046952	839704136	839704136								
Has got a Kinect Sensor?	no	1498408	1300653	8394706								
Uses PC for learning?	no	1	4	3								
How?		1	4	3								
PC usage per day (hours)	0											
Answers from post-test												
Feedback												
Best HCI	1											
Worst HCI	2											

Demographic data [i.e. answers from the pre test]												
Age [10-14]												
Played computer games?	Yes	1	2	3								
Type of games	Ego / Adv. / other (PS Move)	338701732	3387015	25603458								
Has got a Kinect Sensor?	no	95490084	3387015	256034								
Uses PC for learning?	yes	2	6	4								
How?	Learning Games	7	6	4								
PC usage per day (hours)	2											
Answers from post-test												
Feedback												
Best HCI	3											
Worst HCI	2											

Test person	HCI	13						2						3						
		"Learning"			"Walking-In-Place"			"Distance-To-Velocity"			"Learning"			"Walking-In-Place"			"Distance-To-Velocity"			
C	E	#E	E%	C	E	#E	E%	C	E	#E	E%	C	E	#E	E%	C	E	#E	E%	
98.7	0.0	0	0.0	92.6	0.0	0	0.0	78.7	0.0	0	0.0	8.0	0.0	0	0.0	8.0	0.0	0	0.0	
R	R-N	Hits	%	R	R-N	Hits	%	R	R-N	Hits	%	R	R-N	Hits	%	R	R-N	Hits	%	
10	50	3	0.200	6	30	2	0.133	4	20	2	0.133	4	20	1	0.067	4	20	1	0.067	
Mental Demand																				
Physical Demand																				
Temporal Demand																				
Performance																				
Effort	15	7.5	0.333	11	5.5	4	0.267	2	10	3	0.200	5	4	2.667	5	4	2.667	5	4	2.667
Frustration	10	50	4	0.267	3	15	5	0.333	1	5	5	0.333	1	5	5	0.333	1	5	5	0.333
Overall	58.6667																			
Overall	31.3333																			
Overall	9.0000																			

Test person	HCI	14						2						3						
		"Learning"			"Walking-In-Place"			"Distance-To-Velocity"			"Learning"			"Walking-In-Place"			"Distance-To-Velocity"			
C	E	#E	E%	C	E	#E	E%	C	E	#E	E%	C	E	#E	E%	C	E	#E	E%	
16.5	8.5	2	7.3	14.82	34.8	2	23.3	73.5	8.4	4	11.4	11.4	0.0	0	0.0	11.4	0.0	0	0.0	
R	R-N	Hits	%	R	R-N	Hits	%	R	R-N	Hits	%	R	R-N	Hits	%	R	R-N	Hits	%	
8	40	2	0.133	14	70	4	0.267	13	65	4	0.267	13	65	4	0.267	13	65	4	0.267	
Mental Demand																				
Physical Demand																				
Temporal Demand																				
Performance																				
Effort	14	70	0	0.000	14	70	3	0.200	13	65	4	0.267	13	65	4	0.267	13	65	4	0.267
Frustration	17	85	4	0.267	17	85	5	0.333	12	60	0	0.000	12	60	0	0.000	12	60	0	0.000
Overall	60.6667																			
Overall	75.3333																			
Overall	67.3333																			

Test person	HCI	15						2						3						
		"Learning"			"Walking-In-Place"			"Distance-To-Velocity"			"Learning"			"Walking-In-Place"			"Distance-To-Velocity"			
C	E	#E	E%	C	E	#E	E%	C	E	#E	E%	C	E	#E	E%	C	E	#E	E%	
14.7	49.0	18	34.6	125.6	35.4	3	20.2	159.4	119.0	15	74.7	15.0	0.0	0	0.0	15.0	0.0	0	0.0	
R	R-N	Hits	%	R	R-N	Hits	%	R	R-N	Hits	%	R	R-N	Hits	%	R	R-N	Hits	%	
11	55	3	0.200	16	80	2	0.133	11	55	2	0.133	11	55	2	0.133	11	55	2	0.133	
Mental Demand																				
Physical Demand																				
Temporal Demand																				
Performance																				
Effort	11	55	5	0.333	11	55	5	0.333	3	15	5	0.333	3	15	5	0.333	3	15	5	0.333
Frustration	11	55	3	0.200	11	55	2	0.133	1	5	3	0.200	1	5	3	0.200	1	5	3	0.200
Overall	48.3333																			
Overall	68.3333																			
Overall	18.0000																			

Test person	HCI	16						2						3						
		"Learning"			"Walking-In-Place"			"Distance-To-Velocity"			"Learning"			"Walking-In-Place"			"Distance-To-Velocity"			
C	E	#E	E%	C	E	#E	E%	C	E	#E	E%	C	E	#E	E%	C	E	#E	E%	
82.4	34.9	10	42.4	82.2	42.3	10	51.5	90.5	38.7	8	42.8	10.0	0.0	0	0.0	10.0	0.0	0	0.0	
R	R-N	Hits	%	R	R-N	Hits	%	R	R-N	Hits	%	R	R-N	Hits	%	R	R-N	Hits	%	
6	30	4	0.267	20	100	4	0.267	1	5	2	0.133	1	5	2	0.133	1	5	2	0.133	
Mental Demand																				
Physical Demand																				
Temporal Demand																				
Performance																				
Effort	20	100	3	0.200	19	95	2	0.133	1	5	2	0.133	1	5	2	0.133	1	5	2	0.133
Frustration	12	60	1	0.067	20	100	0	0.000	10	50	3	0.200	10	50	3	0.200	10	50	3	0.200
Overall	42.3333																			
Overall	59.3333																			
Overall	14.0000																			





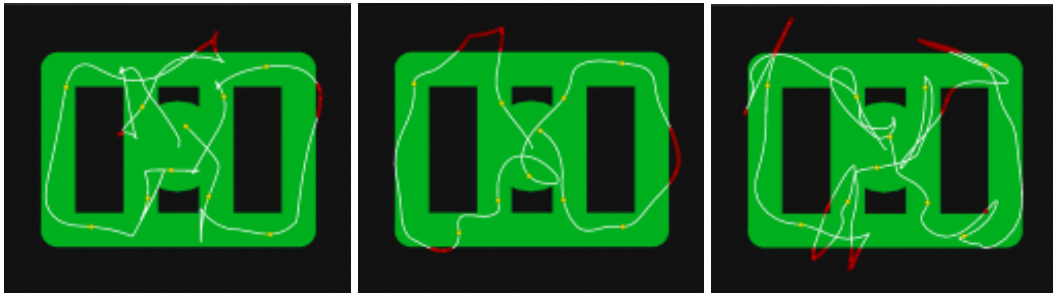
Test person	Z1												♀	3				
	1			2			3			4								
HCI	"Learning"			"Walking-in-Place"			"Distance-To-Velocity"			"Distance-To-Velocity"			3					
	C	E	#E	C	E	#E	C	E	#E	C	E	#E	C	E	#E	C	E	#E
Measured Performance	108.9	38.0	5	16.5	99.0	32.6	10	32.9	126.1	67.9	20	53.8						
NASA TLX																		
Mental Demand	10	50	5	0.333	15	75	5	0.333	13	65	5	0.333						
Physical Demand	7	35	2	0.133	15	75	4	0.267	10	50	2	0.133						
Temporal Demand	4	20	3	0.200	3	15	1	0.067	2	10	2	0.133						
Performance	9	45	1	0.067	15	75	1	0.067	16	80	2	0.133						
Effort	10	50	4	0.267	16	80	3	0.200	14	70	4	0.267						
Frustration	1	5	0	0.000	14	70	1	0.067	13	65	0	0.000						
Overall	<b>41,6667</b>			<b>71,6667</b>			<b>59,0000</b>											
Demographic data [i.e. answers from the pre test]													Age [21-40]			Memory_Task		
Played computer games?	no												HCI			1		
Type of games													Original			581328304		
Has got a Kinect sensor?	no												User			264310968		
Uses PC for learning?	yes												#Prefix			583294		
How?	Internet												#Position			3		
PC usage per day (hours)	1															2		
Answers from post-test																2		
Feedback																3		
Best HCI	1															2		
Worst HCI	2															3		
Demographic data [i.e. answers from the pre test]													Age [21-40]			Memory_Task		
Played computer games?	yes												HCI			1		
Type of games	Enter (WII)												Original			978064231		
Has got a Kinect sensor?	no												User			5708317		
Uses PC for learning?	no												#Prefix			2		
How?													#Position			2		
PC usage per day (hours)	>3															7		
Answers from post-test																2		
Feedback																2		
Best HCI	2															2		
Worst HCI	3															3		

Test person	Z2												♀	3				
	1			2			3			4								
HCI	"Learning"			"Walking-in-Place"			"Distance-To-Velocity"			"Distance-To-Velocity"			3					
	C	E	#E	C	E	#E	C	E	#E	C	E	#E	C	E	#E			
Measured Performance	116.7	72.4	16	82.0	85.7	70.6	13	82.1	111.2	61.6	13	55.4						
NASA TLX																		
Mental Demand	15	75	5	0.333	14	70	4	0.267	19	95	5	0.333						
Physical Demand	5	25	1	0.067	6	30	2	0.133	4	20	1	0.067						
Temporal Demand	6	30	0	0.000	4	20	0	0.000	5	25	0	0.000						
Performance	15	75	4	0.267	5	25	5	0.333	15	75	4	0.267						
Effort	13	65	3	0.200	14	70	3	0.200	16	80	2	0.133						
Frustration	9	45	2	0.133	5	25	1	0.067	8	40	3	0.200						
Overall	<b>65,6667</b>			<b>46,6667</b>			<b>71,6667</b>											
Demographic data [i.e. answers from the pre test]													Age [21-40]			Memory_Task		
Played computer games?	yes												HCI			1		
Type of games	Enter (WII)												Original			978064231		
Has got a Kinect sensor?	no												User			5708317		
Uses PC for learning?	no												#Prefix			2		
How?													#Position			2		
PC usage per day (hours)	>3															7		
Answers from post-test																2		
Feedback																2		
Best HCI	2															2		
Worst HCI	3															3		

## A.6. Walk Paths

In the following, the walk paths of the navigation task are listed for all participants. Red lines indicate positions off the path while white lines indicate positions on the path. Yellow dots are landmarks.



Walk paths of participant 1. From left to right: LEAN, WiP, D2V.



Walk paths of participant 2. From left to right: LEAN, WiP, D2V.



Walk paths of participant 3. From left to right: LEAN, WiP, D2V.



Walk paths of participant 4. From left to right: LEAN, WiP, D2V.



Walk paths of participant 5. From left to right: LEAN, WiP, D2V.



Walk paths of participant 6. From left to right: LEAN, WiP, D2V.



Walk paths of participant 7. From left to right: LEAN, WiP, D2V.



Walk paths of participant 8. From left to right: LEAN, WiP, D2V.



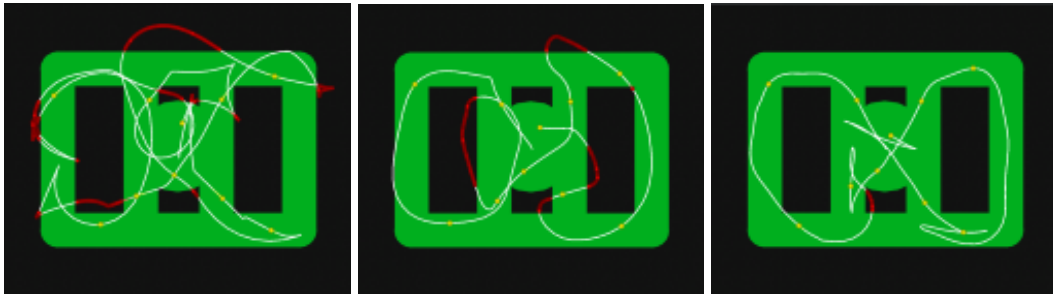
Walk paths of participant 9. From left to right: LEAN, WiP, D2V.



Walk paths of participant 10. From left to right: LEAN, WiP, D2V.



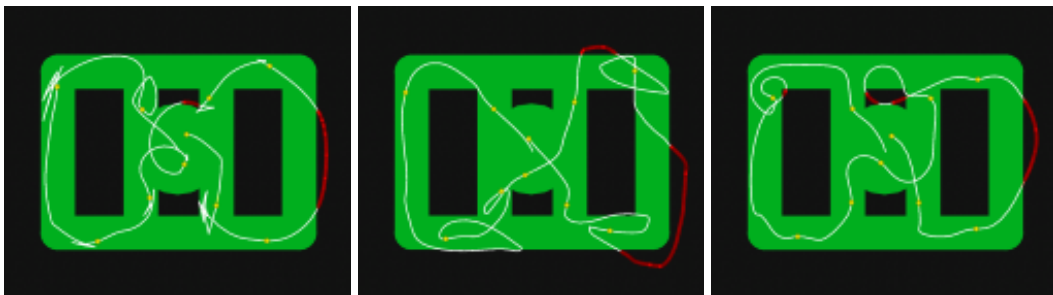
Walk paths of participant 11. From left to right: LEAN, WiP, D2V.



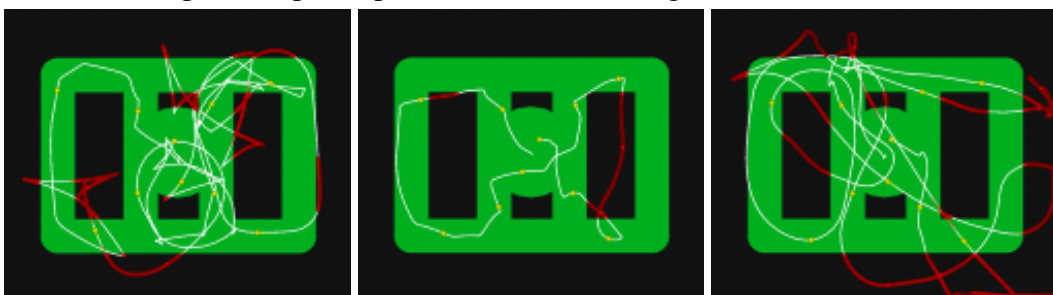
Walk paths of participant 12. From left to right: LEAN, WiP, D2V.



Walk paths of participant 13. From left to right: LEAN, WiP, D2V.



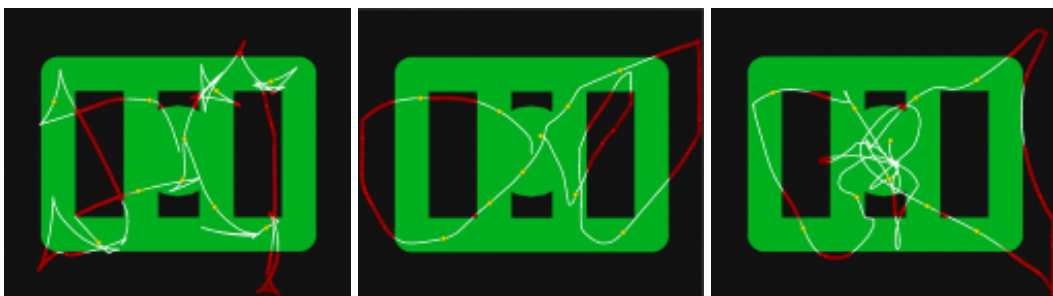
Walk paths of participant 14. From left to right: LEAN, WiP, D2V.



Walk paths of participant 15. From left to right: LEAN, WiP, D2V.



Walk paths of participant 16. From left to right: LEAN, WiP, D2V.



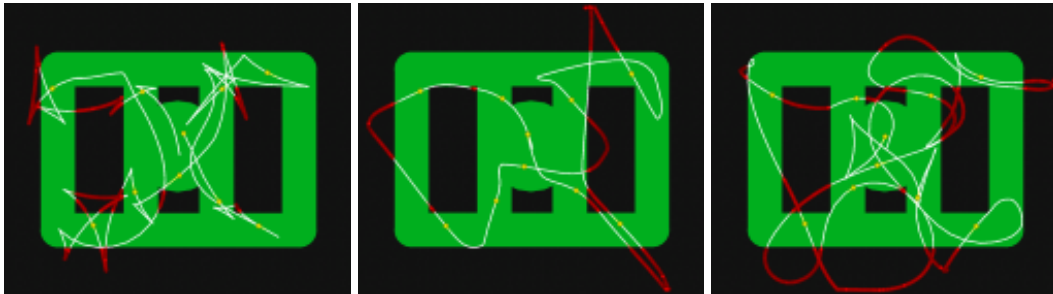
Walk paths of participant 17. From left to right: LEAN, WiP, D2V.



Walk paths of participant 18. From left to right: LEAN, WiP, D2V.



Walk paths of participant 19. From left to right: LEAN, WiP, D2V.



Walk paths of participant 20. From left to right: LEAN, WiP, D2V.



Walk paths of participant 21. From left to right: LEAN, WiP, D2V.



Walk paths of participant 22. From left to right: LEAN, WiP, D2V.

---

## References

- Artino Jr, A. R. (2008). Cognitive load theory and the role of learner experience: An abbreviated review for educational practitioners. *Aace Journal*, 16(4), 425–439.
- Battiste, V., & Bortolussi, M. (1988). Transport pilot workload: A comparison of two subjective techniques. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 32(2), 150-154.
- Bin Tomi, A., & Rambli, D. (2011). A conceptual design for augmented reality games using motion detection as user interface and interaction. In H. Zaman et al. (Eds.), *Visual informatics: Sustaining research and innovations* (Vol. 7067, p. 305-315). Springer Berlin Heidelberg.
- Burdea, G. C., & Coiffet, P. (2003). *Virtual Reality Technology* (2nd ed.). New York, NY, USA: John Wiley & Sons, Inc.
- Byers, J. C., Bittner, A. C., Hill, S. G., Zaklad, A. L., & Christ, R. E. (1988). Workload assessment of a remotely piloted vehicle (rpv) system. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 32(17), 1145-1149.
- Carroll, J. M. (2009). Human–computer interaction. *Encyclopedia of Cognitive Science*.
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and instruction*, 8(4), 293–332.
- Corbett, A. T., Koedinger, K. R., & Anderson, J. R. (1997). Intelligent tutoring systems. *Handbook of humancomputer interaction*, 849–874.
- Crowell III, H. P., Faughn, J. A., Tran, P. K., & Wiley, P. W. (2006). *Improvements in the omni-directional treadmill: Summary report and recommendations for future development* (Tech. Rep.). DTIC Document.
- Dam, P., Braz, P., & Raposo, A. (2013). A study of navigation and selection techniques in virtual environments using microsoft kinect®. In *Virtual augmented and mixed reality. designing and developing augmented and virtual environments* (pp. 139–148). Springer.



- Darken, R. P., Cockayne, W. R., & Carmein, D. (1997). The omni-directional treadmill: A locomotion device for virtual worlds. In *Proceedings of the 10th annual acm symposium on user interface software and technology* (pp. 213–221). New York, NY, USA: ACM.
- Gamma, E., Helm, R., Johnson, R., & Vlissides, J. (1994). *Design patterns: elements of reusable object-oriented software*. Pearson Education.
- Gu, Y. (2014). *SafeChild: Intelligentes Verkehrssicherheitstraining für Kinder in Virtual Reality*. Retrieved 1 August 2014, from <http://safechild.celtech.de/#technology>
- Gu, Y., & Sosnovsky, S. (2014). Recognition of student intentions in a virtual reality training environment. In *Proceedings of the companion publication of the 19th international conference on intelligent user interfaces* (pp. 69–72). New York, NY, USA: ACM.
- Hart, S. G., & Staveland, L. E. (1988). Development of nasa-tlx (task load index): Results of empirical and theoretical research. *Advances in psychology*, 52, 139–183.
- Hart, S. G., & Staveland, L. E. (2002). *NASA Task Load Index*. Retrieved 1 August 2014, from <http://humansystems.arc.nasa.gov/groups/tlx/downloads/TLXScale.pdf>
- ifixit.com. (2010). *Microsoft Kinect Teardown*. Retrieved 30 August 2012, from <http://www.ifixit.com/Teardown/Microsoft-Kinect-Teardown/4066>
- Jacobson, J. (2008). Ancient architecture in virtual reality; does visual immersion really aid learning. *Unpublished doctoral dissertation, School of Information Sciences, University of Pittsburgh*. Retrieved from <http://planetjeff.net/Jacobson2008.pdf>
- kickstarter.com. (2014). *Omni: Move Naturally in Your Favorite Game*. Retrieved 1 September 2014, from <https://s3.amazonaws.com/ksr/projects/476061/photo-main.jpg?1397812032>
- Kovach, S. (2014). *Facebook Buys Oculus VR For \$2 Billion*. Retrieved 5 June 2014, from <http://www.businessinsider.com/facebook-to-buy-oculus-rift-for-2-billion-2014-3>
- Liu, H., & Mihalcea, R. (2007). Of men, women, and computers: Data-driven gender modeling for improved user interfaces. *Proceedings of the International Conference on Weblogs Social Media (ICWSM)*.

- Marsh, W. E., Putnam, M., Kelly, J. W., Dark, V. J., & Oliver, J. H. (2012). The cognitive implications of semi-natural virtual locomotion. In *Virtual reality short papers and posters (vrw), 2012 ieee* (pp. 47–50).
- McKnight, P. E., & Najab, J. (2010). Mann-whitney u test. In *The corsini encyclopedia of psychology*. John Wiley & Sons, Inc.
- Medina, E., Fruland, R., & Weghorst, S. (2008). Virtosphere: Walking in a human size vr 'hamster ball'. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 52(27)*, 2102-2106.
- Microsoft. (2013). *Human Interface Guidelines v1.8*. Retrieved 28 August 2014, from <http://msdn.microsoft.com/en-us/library/jj663791.aspx>
- Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review, 63(2)*, 81.
- Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38(1)*, 63-71.
- Pletcher, T., Bier, K., & Lubitz, D. v. (2000). An immersive virtual reality platform for medical education: Introduction to the medical readiness trainer. In *Proceedings of the 33rd hawaii international conference on system sciences-volume 5 - volume 5* (pp. 5025–). Washington, DC, USA: IEEE Computer Society.
- Poupyrev, I., Tan, D. S., Billingham, M., Kato, H., Regenbrecht, H., & Tetsutani, N. (2002). Developing a generic augmented-reality interface. *Computer, 35(3)*, 44–50.
- PrimeSense. (2010). *The PrimeSensor™ Reference Design 1.08*. Retrieved 25 August 2014, from <http://www.souvr.com/Soft/UploadSoft/201005/2010050617295050.pdf>
- Reid, G. B., & Nygren, T. E. (1988). The subjective workload assessment technique: A scaling procedure for measuring mental workload. *Advances in Psychology, 52*, 185–218.
- Rogers, Y., Sharp, H., & Preece, J. (2007). *Interaction Design: Beyond Human-Computer Interaction* (2nd ed.). Chichester: Wiley.
- Roussos, M., Johnson, A., Leigh, J., Barnes, C., Vasilakis, C., & Moher, T. (1997). The nice project: Narrative, immersive, constructionist/collaborative environments for learning in virtual reality. In *In edmedia/ed-telecom* (pp. 917–922). AACE.

- Rubio, S., Díaz, E., Martín, J., & Puente, J. M. (2004). Evaluation of subjective mental workload: A comparison of swat, nasa-tlx, and workload profile methods. *Applied Psychology*, 53(1), 61–86.
- Shallice, T., McLeod, P., & Lewis, K. (1985). Isolating cognitive modules with the dual-task paradigm: Are speech perception and production separate processes? *The Quarterly Journal of Experimental Psychology Section A*, 37(4), 507-532.
- Slater, M., Usoh, M., & Steed, A. (1995). Taking steps: the influence of a walking technique on presence in virtual reality. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 2(3), 201–219.
- Swapp, D., Williams, J., & Steed, A. (2010, March). The implementation of a novel walking interface within an immersive display. In *3d user interfaces (3dui), 2010 iee symposium on* (p. 71-74).
- van Wyk, E., & de Villiers, R. (2009). Virtual reality training applications for the mining industry. In *Proceedings of the 6th international conference on computer graphics, virtual reality, visualisation and interaction in africa* (pp. 53–63). New York, NY, USA: ACM.
- Viager, M. (2011). *Analysis of Kinect for mobile robots*. Retrieved 1 August 2014, from <http://www.scribd.com/doc/56470872/Analysis-of-Kinect-for-Mobile-Robots-unofficial-Kinect-data-sheet-on-page-27>
- Wikipedia. (2014). *Dance pad*. Retrieved 28 August 2014, from [http://en.wikipedia.org/wiki/Dance\\_pad](http://en.wikipedia.org/wiki/Dance_pad)
- wizdish.com. (2014). *The original patented omni-directional treadmill with no moving parts*. Retrieved from [http://i.ytimg.com/vi/wxbQ\\_XLCN8s/maxresdefault.jpg](http://i.ytimg.com/vi/wxbQ_XLCN8s/maxresdefault.jpg)