Human Media Interaction

UNIVERSITY OF TWENTE.

User experience and User performance of feedback in a Pose Game.

Master Thesis by
Niek J. Hoeijmakers, MSc

Graduation Committee:
Dr. ir. Ronald Poppe
Dr. ir. Dennis Reidsma
Dr. Mannes Poel
Prof. dr. ir. Anton Nijholt

Human Media Interaction,
Department of Electrical Engineering, Mathematics & Computer Science,
University of Twente

December 14th, 2011

Summary

The goal of the research was to provide an enjoyable experience in posing. A pose game was developed that guided players to a desired pose. When the player reached the desired pose his image was placed in a fictional scene as a reward. The game provided feedback in the form of a hint towards the desired pose, an indication of the distance between the desired pose and the player's current pose and, a mirror image of the player on screen. The input for the game was vision based. The Kinect provided the sensor data about the player's pose to the pose game.

This thesis describes the study that aimed at evaluating the feedback of the pose game. Four feedback options were implemented in this study. With each option the hint towards the desired pose was varied. 1) The minimum feedback did not provide a hint. 2) The pose chain feedback showed in succession three silhouettes of poses describing the movement to the desired pose. 3) The inkspot transition had an undefined shape transforming into a silhouette as the player came closer to the pose. And 4) the focus image feedback showed an image of an object indicating the context and the focus of the pose.

The study questioned which of these options had the best user performance, based on the constructs effectiveness and efficiency, and which provided the best user experience, based on the constructs pleasure, engagement and satisfaction. The experiments were done in a laboratory setting and 28 participants, recruited from the University of Twente, provided their experiences with each of the four options through questionnaires and an interview. The order of the feedback was varied over four groups to balance the results.

The results of the survey were analyzed using a repeated measures ANOVA. The feedback options, the participants and the respective measures were within subjects factors and the group of feedback order was the between subjects factor. For the analyses where task completion was a constraint a one-way ANOVA was used with the feedback options, the (un)completed tasks and the respective measure as the factors in the analyses.

The results show that the pose chain feedback option gave the best user performance and provided the best user experience. The reason why most of the participants liked this option the best is that it was the most clear on what needed to be done. The inkspot transition and focus image were favored by some of the participants, because these were more challenging and required more creativity. The minimum feedback was the least favored and in general the participants were the most negative about it.

Preface

First and foremost I dedicate this milestone to my father, Jan Hoeijmakers, who has been with me in spirit all these years.

This thesis marks the end of a long journey for me, as I look back on some fond years of my life. I close my chapter as being an undergraduate student, although I still have a lot more to learn and an interest in a few more courses. I am ready to take the next intersection on the road that is my life. Where this path shall lead me I do not know yet.

I have always had an interest in computer games and interactive entertainment in general. I had a lot of fun experimenting with new input devices in previous research courses. Thus when it came down to choosing the topic of my final project the choice was easy. After a discussion with some of my supervisors I ended up with the project presented in this thesis. Although it started out using a regular webcam at first. This was interesting, because it also presented me with the challenge of learning computer vision and image processing.

I would like to thank my mother, Trudy, and sister, Linda, for encouraging me and standing by me throughout my studies and its detours. I would also like to thank the rest of my family and friends who have given their moral support. All their support has made these years the best experience in my life.

I would like to thank my graduation committee, Ronald Poppe, Dennis Reidsma, Mannes Poel and Anton Nijholt for their guidance and support during all these months. Especially Ronald and Dennis for actively steering me when I was veering off course and slowing me down when I was going too fast through the steps. Also a big thanks to Betsy van Dijk for her advice and help with the statistical analyses. I could always come to you guys with any question I have had. I would also like to thank Lynn Packwood for grammar checking my thesis.

Finally I would like to thank all the volunteers who participated in the experiments of this study. Without them I would not have any data or results for this study.

Thank you all.

Niek Hoeijmakers
December 14th, 2011.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

> **Pose:**
> a way of standing or sitting, especially in order to be photographed, painted, or drawn.
> *Oxford Dictionaries*

Everyone has had his or her picture taken at one time or another. Maybe for the family photo album; for the memories of a nice vacation; or for official documents, such as a passport. The photos for official documents often only contain the person's head, but they still have some strict directives on how a person should pose. Specifications for measurements and facial expression should be adhered to when the picture is taken.

Vacation photos on the other hand have absolutely no directives. The pictures can range from head shots to full body shots. A photo can contain only one person or a very large group. But every person in a picture is captured in a specific pose. These poses vary and can be funny, serious or action shots.

Posing in front of a camera can be fun and the resulting photos can tell a story. The activity of posing in itself can be entertaining, as the television game show "Hole in the Wall" has shown (see Figure 1.1). A game was developed for the research that let players experience the fun of posing.

1.1 The research

The goal of this research was to provide the player with an enjoyable experience in posing. A pose game was developed with the goal to motivate and persuade the player to assume various poses. The feedback of the game should draw the player's attention and persuaded the player to keep on moving until the desired pose was reached. This desired pose was chosen by the game. The feedback provided by the game should steer the player towards the desired pose. Once the player was in the correct pose the game would take a picture and placed the player's image in a fictional scene.

The player needed to pose as a model posed for a photographer. The input was thus noninvasive and relied on vision-based input, such as the Microsoft Kinect Sensor. The pose game could be considered an exertion game. The poses in the game required the use of the entire body. However the player's movements were confined to a small area. The exertion in this game was manipulating your own body into the desired pose and balancing the body to hold the pose. The player needed enough room to assume the strangest poses and to be completely visible in the camera.

Figure 1.1: The game show Hole in the Wall. The participants either successfully assumed the pose carved in the wall that was heading for them, scoring points for their team. Or they were pushed into the water. Source: http://www.realtvaddict.com/2009/10/21/tv-game-shows/

For the pose game that was developed for this study four different feedback methods were implemented. The feedback provided information about the pose that was desired by the game, and information on whether the player's pose approached the desired pose. In all four feedback options there was a background that changed color from red to green when the player's pose came close to the desired pose. In addition, all the feedback options displayed the player's image like a reflection in a mirror. The difference between the feedback options was the hint they provided to the desired pose (see Chapter 3):

Minimum feedback provided no hint to the desired pose.

Pose chain feedback provided three silhouettes in succession leading to the desired pose.

Inkspot transition feedback provided an abstract shape that became more concrete and transitioned into the silhouette of the desired pose as the player approached it.

Focus image provided the image of an object whose context was an indication to the desired pose. Its position indicated the focus of the pose.

1.2 The user study

The focus of the user study was to evaluate how the four feedback feedback options that the pose game provided affected the user experience and the user performance. For the study two main research questions were formulated:

1. Which of the feedback options provides the best user experience?
2. Which of the feedback options has the best user performance?

User experience is still a very broad concept without a clear definition or decent measurement [Has05], unlike user performance which is well defined. In this study user experience was narrowed down to the constructs *pleasure*, *engagement* and *satisfaction*. User performance was measured by *effectiveness* and *efficiency*.

The participants in the study experienced each feedback option once where they had to find the desired pose. The setup used a Kinect sensor as the input for the participants and a large display for the output of the feedback. The measurements of the participant's performance were logged by the game. The participants filled out a questionnaire after each task and participated in an interview at the end of the study.

1.3 Overview of this report

This report will further detail the research and the results. *User performance*, *user experience* and *exertion games* are central terms to this study. In chapter 2 these terms will be discussed in more detail.

The next two chapters will discuss the design and setup of the experiment. In chapter 3 a detailed explanation of the assignment will be provided. While in chapter 4 the methodology will be discussed, starting with the formulation of the research questions and subquestions. Next the experimental design, setup and procedure will be discussed. Followed by a description of the expected participants. Finally in this chapter the hypotheses for the research will be discussed.

A demographic profile and the data gathered will be discussed and analyzed in chapter 5. Remarkable results will be highlighted and a general answer to the research questions will be provided based on the results. In chapter 6 we will take a critical view on the results.

In chapter 7 we will summarize the experiment and the concluding answers to the research questions will be provided. Finally in chapter 8 recommendations for future research will be given.

Chapter 2

Related work

In this study *user experience* and *user performance* were two central subjects to be evaluated with an *exertion game*. Where user performance is well defined, user experience is still lacking. This chapter will discuss user performance first, user experience second and finally what an exertion game is.

2.1 User performance

User performance is an objective term. Its measure is based on the two objective constructs *effectiveness* and *efficiency*. User performance is well defined by the 'International Organization for Standards' (ISO). They define user performance as the accuracy and completeness with which a user can achieve their goals using the system and the resources required to achieve the goals [ISO98].

To put it simply. User performance looks at whether the product does what its purpose is; how well the product does it; and how much it costs to complete its purpose. User performance can also be used to quantify and objectively compare several similar products, such as exertion interfaces [FHVN09].

Measuring user performance can been done relatively easily and objectively, considering the definition [ISO98]. For example Manresa-Yee et al. were developing a vision based interface for disabled people. They assessed the performance of their product by having six subjects use their product in several sessions [MYPVP10]. The researchers noted whether the subjects were successful in their assigned task. How long the subjects needed to complete the tasks. And how fatigued they were after the task. The researchers compared the results with the performance of products that the subjects usually used.

Basically to measure performance is to measure how often someone fails or has success using the product. How close the user gets to success. The costs of a product can be several things such as time [FHVN09] or physical effort [MYPVP10]. The efficiency can be measured over a single task, over several tasks or the number of task that can be done within a set amount of resources.

2.2 User experience

The definition of user experience is not as well defined in contrast with user performance. User experience is a very broad subject. Intuitively we may understand what is meant by user experience. However Hassenzahl mentions that we are still far from a coherent understanding [Has05]. And Wright et al. state that "Experience is an elusive concept that resists specification and finalisation" [WMM05]. Yet, Manresa-Yee et al. argue that involving users and measuring user experience is essential in each stage of development [MYPVP10]. The questions remain 'what is user experience?' and 'how can we measure user experience?'

2.2.1 Defining user experience

Some of the terms that immediately come to mind when looking at user experience are enjoyment [BFH05], pleasure [MR74], satisfaction [DTGP10] and being engaged [BFC+09]. However, none of these terms seem to fully cover user experience and are but a few of a wide range of constructs to measure. For most of these subjects there is no consensus on the definition either. And yet, user experience has become a catchphrase that everyone acknowledges.

ISO only considers *satisfaction* and provides a definition as "freedom from discomfort, and positive attitudes towards the use of the product" [ISO98]. However other researchers have given other definitions of satisfaction in their works.

Deng et al. adopt the definition from Oliver: "Satisfaction is the consumer's fulfillment response. It is a judgment that a product or service provided a pleasurable level of consumption-related fulfillment, including levels of under- or over-fulfillment" [DTGP10]. In their view pleasure is a sub-construct of satisfaction. Their definition, however, contradicts the definition of Blythe and Hassenzahl. According to them satisfaction and pleasure are two equal constructs with an opposing aspect [BH05]. Satisfaction is the result coming from the confirmation of expectations, while pleasure is the result of a deviation from expectations.

Blythe and Hassenzahl also argue the difference between *fun* and *pleasure*. A fun experience is determined by its value as a distraction and being trivial, repetitive, a spectacle and a transgression. While pleasure is the opposite and determined by its value as an absorption, being relevant, progressive, aesthetic and being committed to the activity [BH05]. Fun is a short experience, while pleasure can take up a long time. To Blythe and Hassenzahl work can be pleasurable but not fun.

Sengers sees an experience of fun in a similar fashion, however does not make the explicit distinction between fun and pleasure [Sen05]. She argues that due to Taylorism our everyday lives and especially work have become a very bland and repetitive experience. Thus for fun we spent our unplanned time doing activities, which maximize pleasure and minimize task achievement.

When discussing experience, *flow* is an often recurring term. The definition of flow that is most used in research was introduced by Csikszentmihalyi's flow theory [MC96]. People experience flow when the activity they are involved in has a subtle balance between challenge and skill [BH05]. Brockmyer et al. consider flow to be a stage of 'engagement' along with immersion, presence and absorption [BFC+09]. Brockmyer et al. see the construct *engagement* as a generic indicator of game involvement.

For designing experiences Sengers argues that social and cultural aspects have to be considered as well [Sen05]. Hassenzahl even comments that an experience is always mediated by the situation of the user [Has05]. A designer or developer can build a system to have a certain product character and is targeted to provide a specific experience. However the result as to whether the user actually experiences the intended experience is painted by the situation of the user. The situation is determined by the larger social and cultural background [Sen05] and by the current mindset and personal history of the user [Has05]. Blythe and Hassenzahl provide as an example the way in which the definition of *fun* has changed through the course of history due to social and cultural changes [BH05].

Experience is the result of the situation and the apparent product character [Has05]. Fogg argues in [Fog03] that people often respond better to technology when there is a common element between the user and the product. He also argues that users experience the use of a product better when a positive psyche can be inferred from the product. The psyche and the product's function need to be compatible with the user's situation. Thus where user performance can be attributed to the product, user experience is attributed primarily to the user.

2.2.2 Measuring user experience

Measuring user experience is more difficult than measuring user performance, due to the vague definition and subjective nature of user experience [WMM05]. There is no method that can easily measure user experience entirely. Any measure created only covers a part. Thus for measuring it is better to focus on a construct or constructs that are the most meaningful to one's own product/project. Various researchers have developed frameworks for measuring experience by focusing on a number of constructs [HDG10, WMM05, Shi09, DTGP10]. Some studies offer actual questionnaires that measure a certain construct of experience [BFC$^+$09, PTF$^+$95, BL94].

Moneta and Csikszentmihalyi mention that "the optimal approach to studying experience is the collection of repeated measures" [MC96]. This means that users reported on their experience at different times and/or on different occasions. They used the Experience Sampling Method to measure the constructs concentration, the wish to do the activity, involvement and happiness. They attempted to find a correlation with the predictors *perceived challenge* and *perceived skill*.

Taylor and Agamanolis studied the satisfaction of a standard telecare product by asking users about their behavior in wearing the product [TA10]. While Manresa-Yee et al. asked their subjects about how satisfied they were [MYPVP10]. Deng et al. determined satisfaction based on cognitive absorption as a predictor in their model [DTGP10], which eventually determines the intention to continue to use the product.

Absorption was also measured in the work of Brockmyer et al. In their work they focussed on the potential for people to get engaged in video game-playing [BFC$^+$09]. Their Game Engagement Questionnaire had 19 questions each measuring one of the four different stages of engagement they considered (immersion, presence, flow and absorption).

Some measures of experience are based on emotions. Mehrabian and Russell devised a questionnaire to measure *pleasure*, *arousal* and *dominance* based on 18 opposing emotional adjectives, they called the Semantic Differential Scale [MR74]. Bradley and Lang later verified and simplified this scale into their Self-Assessment Manikin [BL94], with only three items represented with pictures. Hassenzahl et al. developed a framework where needs (competence, relatedness, popularity and more) and affects (from PANAS: afraid, scared, excited, inspired and more) are predictors for pragmatic and hedonic qualities of user experience [HDG10].

Other frameworks and questionnaires approach the experience from a motivation direction. Pelletier et al. measured a person's experience in sports by determining their motivation [PTF$^+$95]. The questionnaire was focussed on determining intrinsic motivation, extrinsic motivation and amotivation. Intrinsic and extrinsic motivation were also a part of the model developed by Shin, who used a modified Technology Acceptance Model [Shi09].

2.3 Exertion games

In the past decade there has been a rise in exertion gaming with the release of cheap commercially available motion sensing devices, such as the Sony Eye Toy[1], Nintendo Wii[2] and Microsoft Kinect[3]. One of the most successful games is Dance Dance Revolution (DDR)[4]. It is one of the earliest commercial successes and has made its way into a number of scientific works [And07, TP10]. Although attempts have been made since the 1980's [NvDR08].

The idea behind an exertion game is that physical exertion is part of the gameplay experience and gameplay is part of the exertion experience [NvDR08]. Often the exertion is used as input, either through a motion sensing device or environment [NPvD+11]. But there are projects where exertion is the result of playing a game [SMM+11].

Scientific research in exertion games goes from developing and assessing new input devices [BTN04, FHVN09, LMW+11] to evaluating the effects of playing exertion games [MST+07, RN08, TP10]. With the latter it is mostly about evaluating physiological and psychological responses. The psychological effects can further be divided into emotional effects, such as mood, and behavioral effects, such as social interaction.

2.3.1 Effects of exertion games

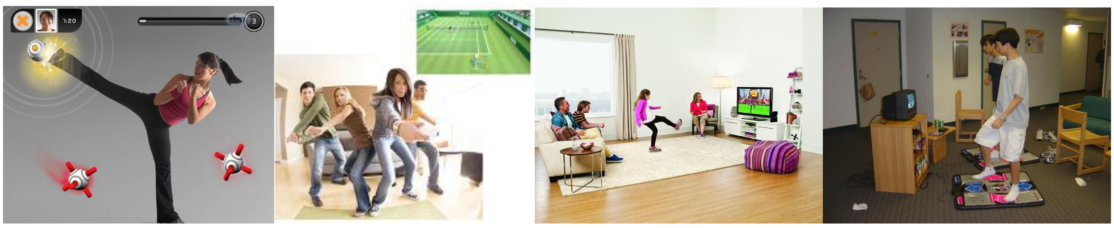
When Nintendo first introduced the Wii console its arguments for motion control were that it makes gaming more enjoyable and gets gamers to be more active. With the introduction of the game Wii Fit Nintendo wanted to show the health benefits of gaming. The scientific community responded by evaluating such a claim and looking at how to

[1] `http://blog.us.playstation.com/2010/11/03/eyetoy-innovation-and-beyond/`, last viewed 14-09-2011.

[2] `http://www.nintendo.com/wii`, last viewed 14-09-2011.

[3] `www.xbox.com/kinect`, last viewed 14-09-2011.

[4] `http://www.ddrfreak.com/`, last viewed 14-09-2011.



(a) Sony Eye Toy (b) Nintendo Wii (c) Microsoft Kinect (d) Dance Dance Revolution

Figure 2.1: Commercial successes in exertion gaming.
(a) A screen capture from the game Eye Toy Kinetic (source: `http://www.dignews.com/platforms/ps2/ps2-reviews/eye-toy-kinetic-review/`, last viewed 21-9-2011).
(b) Nintendo Wii promotional image of players holding a Wii remote (source: `http://damyn.wordpress.com/2010/03/16/concept-van-nu/`, last viewed 21-9-2011).
(d) Kinect promotional image of people playing a Kinect game (source: `http://www.gamestar.hu/jatek/microsoft-kinect.html`, last viewed 21-9-2011).
(d) Two people playing DDR with a dance mat (source: `http://ballroomdancess.com/2011/08/20/dance-revolution/`, last viewed 21-9-2011).

utilize it in medical contexts [HV09, TP10, GGLAC11].

Obesity and physical development in children and obesity in general has always been of great concern to parents and physicians [HHTR05]. Which is only made worse due to the sedentary lifestyles developed with the growing popularity of videogames [PB10, BDDM10]. In 2007 Mark Anders reported on a study done by the *American Council on Exercise* where the exertion video game DDR was evaluated as a viable alternative for physical exercise in schools [And07]. The results showed that the energy expenditure on the higher levels were sufficient for the recommended daily physical activity.

Thin and Poole performed a similar study, where they evaluated the health benefits of DDR and two Wii Fit games [TP10]. Their conclusion was that the specific games were not sufficient for the recommended daily physical activity. However they commented that this might be due to the subjects' inexperience with the game. This suggestion had been corroborated by the study performed by Sell et al. Their results showed that the energy expenditure with experienced DDR players was higher than the inexperienced players [SLT08].

A number of studies focussed on the physiological benefits of exertion games, each with their own target demographic among the healthy population and using self-made or existing commercial games [PB10, BDDM10, WRK11]. Other studies focussed on helping disabled people exercise their muscles [SMM$^+$11] or to even rehabilitate after being seriously injured [GGLAC11]. Manresa-Yee et al. developed a vision-based control interface for people with cerebral palsy. Originally this interface was meant to provide easier access to computers, however as a side effect it strengthened the subjects' neck muscles and improved control over head movements [MYPVP10].

As the above mentioned studies focussed on the physiological effects, there were a few that actually looked at the psychological effects. Some of these studies looked at emotional effects such as mood changes [RN08], while others considered the behavioral effects, such as social interaction and bonding [MGVA08]. A reoccurring measure in almost every study, even with the former studies, was whether the subjects enjoyed the exertion game. And often the interactive videogame prevailed in enjoyment over the sedentary counterpart or a regular fitness exercise [PB10, SLT08, TP10, BDDM10]. However this might also be part of the novelty that the subjects were experiencing with exertion games [FHVN09].

Exertion games had a positive effect on a person's mood. More so than regular videogames. It was the exercise part of the video game that had this emotional effect [RN08]. Russell and Newton's study showed no difference in mood effect between a regular exercise and an interactive videogame. But a large difference between the interactive and the sedentary videogame. Mueller and colleagues also encountered this effect in their results. In their *Table Tennis for Three* project three subjects played together. The pleasure and fun of playing skipped over into the interviews afterwards as they were more vivid than usual [MGVA08]. Exertion did not only result in stronger positive moods. When a player was confronted with violent and aggressive stimulants during game play, the more aggressive mindset seemed to last longer as well [MDHS10].

The focus of Mueller's projects was not to measure a person's individual mood, but to measure the behavior of social interaction between people. The results from projects, such as *Breakout for Two*, showed that exertion games facilitated the social bonding process between individuals [MST$^+$07]. They also discovered that players tended to be more competitive towards each other even after the match was over [MBB10]. The social interaction was also the main design goal for the *Age Invader* project. However Khoo et al. designed it to facilitate social interaction between various generations within a family [KMC+07].

2.3.2 The technology behind exertion games

At the other end is the development of the technology, both as hardware and software. The technology can be simple or very complex. The main criteria is that it facilitates physical interaction between all the actors [NvDR08].

The Age Invaders project required a room with a tiled floor containing RFID sensors, RFID tagged slippers and a large display to show the virtual environment for the interaction between grandparents and grandchildren, while the parents could join over the internet via the computer [KMC+07]. The physical interaction took place over a larger area. The input was given to the computer with the RFID tags. The Breakout for Two project also used a large room as players need to shot a ball against the wall [MST+07]. However it was the impact of the ball against the wall that was the input to the virtual environment. For the social aspect a video conference image was projected on the wall.

Not all exertion games required a lot of space for people to move around in. Li et al. implemented a *whack-a-mole* game that players could control by swinging a brightly colored hammer around in front of a camera [LCW+08]. And Mueller and colleagues studied social interaction facilitated by playing on bongos in a game of *Donkey Konga* [MBB10]. This type of interaction only relied on physical activity of the upper body, while the feet were stationary.

However limiting the space to move in does not limit the exertion to just the upper body. Even in a limited space the entire body can be used. For example by strapping the player in a harness [MST+07] or choosing an activity that does not require much space, such as rap dancing [RvWP+06] or a simple fitness exercise [RZvWR06]. In *FlyGuy* Mueller and colleagues strapped the player in a harness. By twisting his body and using gravity the player could experience a virtual flight in a hang glider.

With the virtual rap dancer the player had to step on a DDR dance mat to start the interaction [RNRH05]. A camera then recorded and analyzed the player's moves in order for the virtual rap dancer to dance along. This dancer provided feedback by either following the player or even leading the player by teaching new dance moves. The interaction only lasted as long as the player had one foot on the mat [FHVN09].

A great number of input devices have been developed. Some of the commercial devices were even considered for serious applications [dHGP08, Fik10]. The devices could be simple and developed for a single game. Or devices could be developed for a more general range of applications, such as the Wii Remote, Balance Board, Kinect or Joyfoot [FHVN09, Sol11, BTN04]. What they should do is provide the necessary information about the exertion to the virtual environment, such as the impact of a ball against a wall [MST+07].

The data from the devices needs to be interpreted by the game and converted into the right manipulation. The press of a button is often mapped to a single action, as with a dance mat and bongos [RNRH05, MBB10]. A device such as the Wii Remote uses accelerometers and a gyroscope [FHVN09]. Foot operated devices, such as the Joyfoot and Wii Balance Board, use pressure sensors to determine the player's weight distribution [BTN04, dHGP08]. Cameras, such as the Kinect, are the least invasive devices but they give the most ambiguous of input. In a camera image the player first needs to be recognized [RSH+05, LMW+11] before the orientation and movement of the player can be determined [BD01].

Nijholt et al. discuss in [NvDR08] some options of measuring experience in exertion games. Besides questionnaires, camera's and other sensors can help us observe a subject in the near future. One step further is to measure physiological information through devices such as heart rate monitors. However Dix et al. argue that the more invasive a measure is the more it influences the measurements [DFAB03]. The same can be said about input devices. The invasiveness of a device can alter the experience of the exertion game. As in simulating a guitar by using either a toy guitar or an air guitar [MBB10].

Chapter 3

The master's thesis assignment

This chapter will explain what the master's thesis assignment was. The assignment was to create a posing game using a vision-based interface, the Microsoft Kinect sensor, that could provide a good experience in playing. The aim of the research was to develop an entertaining and enjoyable experience of posing in front of a camera for players. This study was to evaluate the user experience and user performance of the immediate feedback of the pose game during gameplay.

This chapter will deal with the research that this study was a part of and the study's focus. The four feedback options that were implemented for this study will be presented. Finally the evaluation in this study will be discussed.

3.1 The project

The general project was to create a game that would engage players into assuming a specific pose in front of a camera in an enjoyable and entertaining manner. A player would step inside a booth, where he would see the game displayed on a large screen. Inside the booth the player would be motivated by the game to position himself in the pose. It was up to the player to find out what pose the game wanted. The game provided feedback to the desired pose and how far the player was still removed from the pose.

The pose game continued as long as the player did not find all the poses. The player would lose the game when he would forfeit and walked away. When a pose was found the pose game took a picture of the player in the desired pose and put him in a different (fictional) scene. A new background, foreground and props would be added to place the player in a different setting and give a (new) meaning to the pose (see figure 3.1 for examples).

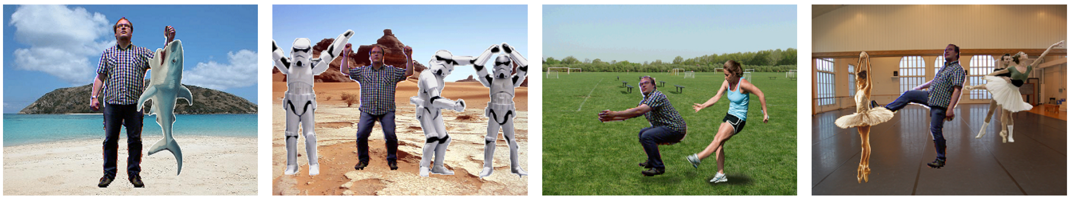


Figure 3.1: These pictures are examples of rewards created by the pose game. The player was placed in a scene with varying background and foreground.

The interaction of the game was considered in three stages. The first stage was attracting the player's attention in that he was dealing with an interactive game and something could be won if he would complete all the tasks.

The second stage was engaging the player to continue playing. The game would provide immediate feedback based on the player's actions. This feedback guided the player to the desired pose that the game selected. This stage was what this study focussed on and had several options implemented (see Section 3.2).

The last stage was rewarding the player for completing a single or all the tasks. This stage involved notifying the player of a task completion and the completion of the game. For completing the game the player was provided with the snapshots that the game took of the player in the desired poses.

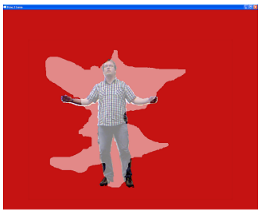
3.2 The feedback options in the study

The focus was the immediate feedback that the player got when he was playing the game. The feedback showed the player's progress towards the completion of a single task and a hint on which pose the game desired of the player. In addition, the feedback also needed to stimulate the player to continue playing until he completed one or more tasks in order to keep the game going. The focus of this study was to evaluate the user experience and user performance of each feedback.

Four different feedback options were designed and implemented. All options used only a visual modality for presenting the necessary feedback information to the player. This lowered the number of variables that could influence the results and kept the study to a manageable size. The aesthetics were kept simple. Fogg mentioned that the visual appeal had an effect on how people perceive technology. When the subject found the game appealing he would be more cooperative and more positive in his ratings and when it was not appealing the subject would be more negative. The appeal was however dependent on each person and their personality [Fog03].

The graphics would contain nothing more than was required to provide the necessary feedback. The feedback had three gameplay elements, of which two were consistent with all four feedback options. The first element gave information about the player himself. This was realized by displaying the image captured by the Kinect sensor on screen (see

(a) Visuals of the pose chain feedback



(b) Visuals of the inkspot transition feedback

(c) Visuals of the focus image feedback

(d) Visuals of the minimum feedback

Figure 3.2: These show the visuals of each feedback option. Each feedback option featured a color changing background and the player's image on screen. **(A) Pose chain:** added the silhouette for the player to copy. **(B) Inkspot transition:** had an undefined shape (the inkspot) that transformed to a silhouette of the pose. **(C) Focus Image:** featured an image indicating the most important aspect and a possible context of the pose. **(D) Minimum:** had no additional visuals to its feedback.

figure 3.2). The image of the player was mirrored.

The second element gave an indication on whether the player got closer to the desired pose or not. This element was realized by having the background color transition back and forth between two colors. The background became more red when the distance increased and more green when the distance decreased. The last element of the gameplay gave a hint on the desired pose. This hint was different for each feedback option.

3.2.1 Minimum feedback

The first feedback option did not give a hint to what pose was desired by the game. This option, shown in figure 3.2d, only included the first two gameplay elements common to all the feedback options. It excluded the third gameplay element.

With this feedback the player only had the background color that transitioned between red and green to work with. Like a game of *warm and cold* the player had to search for the correct pose, only to get an idea of the desired pose when he reached it.

This feedback option also formed the basis of the other three options.

3.2.2 Pose chain feedback

This option was inspired by the *'Hole in the Wall'* (see figure 1.1) concept and is shown in figure 3.2a. The hint was given by displaying the silhouette of a person performing the target pose. The player only needed to align himself with the silhouette and copy this pose.

When the player successfully copied the pose the game went to the next pose in the chain. The final pose in the chain was also the desired pose that would produce the reward. The complete sequence of poses (three in total) described the movement of the body from a base stance towards the desired pose.

3.2.3 Inkspot transition feedback

This option, shown in figure 3.2b, gave the hint on the pose in the form of an inkspot. The inkspot started out as a meaningless shape and would shrink towards a silhouette of a person performing the desired pose. The inkspot would only shrink when the player got closer to the pose he needed to assume. The inkspot would not increase in size when the distance increased again.

In this feedback the player needed to explore for a clue to the desired pose, rewarding the player with a smaller and more defined shape when he got closer to his goal. After the player came close enough the shape would be an example silhouette for the player to copy, like in the previous feedback but only the last pose in the chain.

3.2.4 Focus image feedback

In this feedback option the hint was given in the form of an image, as shown in figure 3.2c. The image provided a meaningful context to the pose the player needed to perform. It also provided the area of focus in the screen where the player had to perform the pose.

In this feedback the player had nothing to copy and was only given an assignment on what to do. Examples of images were a chair for the player to sit on and an umbrella for the player to hold.

3.3 The evaluation in the study

In chapter 2 the terms user performance, user experience and exertion games were explored. These were most central to this study and influenced the experimental design in chapter 4. Posing is, by definition, standing or sitting and is not as physically active as dancing, walking or waving. Still in this game the player constantly needed to move his body or limbs to find the desired pose, which made the pose game an exertion game.

Unlike many studies involving exertion games this study did not focus on the exertion part of the game and how the exertion was input in the game. The study evaluated user performance and user experience of the feedback in an exertion game. User performance is an objective concept that is well defined by the International Organization of Standards, as is discussed in section 2.1 [ISO98]. It was measured by how adequately the feedback of the pose game could fulfill its purpose, which was to guide the player to the correct pose that was desired by the game.

User experience on the other hand is a subjective concept, as is discussed in section 2.2. Even though many have an understanding of this concept its definition is still very broad, not coherent and rather elusive to finalize [Has05, WMM05]. Nor is there a measure that can adequately measure user experience in general. As concluded in section 2.2 it is better to select some sub-concepts fitting to the study. For an experience the immediate feedback should have made playing pleasurable and have gotten the player engaged in posing. In addition the player should be satisfied with the feedback he got from the game.

Studies have indicated that exertion could have an influence on the mental state of a person, if the intensity was high enough. However it was not expected that the player would put in much exertion. Even though the full body was used in this game the player's space to move was limited to the range of the Kinect sensor's field of view, which was used as a noninvasive means of registering player input. The player did not need to attach visual markers nor hold the weight of a prop for visual recognition during gameplay. Studies also showed that the intensity of the exertion increased as the player was more familiar and skilled with the game. The participants in this study were not familiar with the game and its feedback, even though the activity of posing might be simple. There was no time limit or another constraint that forced the participant to hasten in finding the desired pose. The participants could play the game at their own leisure and stop when they chose to.

The intensity of the physical exertion was expected to be more than that of a sedentary game. However the difference was small enough that psychological influences of the exertion on the results were expected to be negligible.

Chapter 4

Methodology

This study was a user evaluation in a laboratory setting with the goal of evaluating the user experience and user performance with the four feedback options of the pose game (see Chapter 3). The study measured the pleasure, engagement and satisfaction constructs for user experience and effectiveness and efficiency constructs for user performance. The survey consisted mostly of questions asked in a written questionnaire before and during the experiment and an oral interview after the experiment. These were aided by observations and computer logged data.

4.1 Research questions

By evaluating *user experience* the study questioned the participants about their experience with each feedback option and distilled which option had the best user experience. The overall experience was assessed through several constructs as user experience is still a very broad term. These constructs were pleasure, engagement and satisfaction of the feedback elements (see Subsection 4.2.2).

In addition this study looked at the user performance of the feedback options. One feedback might move the participant to the desired pose more easier than others, while another would not get the participant to assume the pose at all. The user performance was measured by the constructs of effectiveness and efficiency (see Subsection 4.2.3).

For this study the following research questions and subquestions were formulated:

1. *Which of the feedback options provides the best user experience?*
   (a) Which of the feedback options provides the most pleasure?
   (b) Which of the feedback options is the most engaging?
   (c) With which of the feedback options are the participants most satisfied?
2. *Which of the feedback options has the best user performance?*
   (a) Which feedback option is the most effective?
   (b) Which feedback option is the most efficient?

Experience and performance with feedback can be contributed to both the gameplay (functionality) and aesthetics (graphics, sound) [Fog03]. The study focussed on the gameplay.

(a) Umbrella (b) Weightlifting (c) Sit (d) Kick

Figure 4.1: These were the four poses used in the experiment that the participants had to assume.

4.2 Experiment design

This section will discuss the design of the experiment. It will discuss the independent variables, which were implemented in the pose game. And the dependent variables of user experience and user performance, that were measured using four techniques for data collection [DFAB03]. First a written questionnaire before and during the experiment, second an oral interview that was recorded on audio, third computer logging implemented in the pose game and notes taken by the experimenter was the final technique.

4.2.1 Independent variables

The independent variable that this study was interested in was the feedback. The study was set up as a within-subjects experiment that had each participant work with each of the feedback options. As the order in which the options were presented could have an influence on the results, the participants were divided into four groups. The order of the options was varied by applying a Latin's square.

Group #1 Pose chain → Inkspot transition → Focus image → Minimum.
Group #2 Inkspot transition → Minimum → Pose chain → Focus image.
Group #3 Focus image → Pose chain → Minimum → Inkspot transition.
Group #4 Minimum → Focus image → Inkspot transition → Pose chain.

Another independent variable was the set of poses. Repeated execution of the same pose would influence the results due to a learning effect. Thus four different poses were selected to be executed by the participants. The differences between poses could have an effect on how the feedback would be perceived. A fixed order, as presented in figure 4.1, was used to make sure that the use of each pose was equally spread over the feedback options. The final independent variable was the participants, which will be discussed in Section 4.5.

4.2.2 Dependent variables of user experience

User experience was one of the two overall dependent variables. However it is still an ambiguous and very broad concept [Has05]. User experience was further specified into the dependent variables *pleasure*, *engagement* and *satisfaction*. These variables were measured with several items in a questionnaire and an interview (see Appendix A).

Pleasure

Pleasure was measured by six bipolar adjective pairs for each feedback. These items were based on the Semantic Differential Scale originally devised by Mehrabian and

Russel [MR74] and later adapted by Bradley and Lang into the Self-Assessment Manikin [BL94]. Bradley and Lang verified that six of the 18 items represented the construct pleasure.

The Self-Assessment Manikin adapted the bipolar adjectives pairs into a single series of pictures. However some text was still needed to make sure everyone understood the context. Thus Mehrabian's six bipolar adjective pairs for pleasure were used to measure pleasure in this study. The items were scored on a 9-point Likert scale.

Engagement

Brockmyer et al. devised, tested and verified a set of items to measure a persons potential to become engaged in video games, called the Game Engagement Questionnaire [BFC+09]. They also used it to measure how engaged subjects were after playing the videogame 'S.T.A.L.K.E.R.'. An adapted version of Brockmyer et al. game engagement questionnaire was used to measure engagement in this study. Even though their questionnaire was based on a sedentary game, it was not expected that the exertion in this study would be of the intensity that it would influence the results [DFAB03, NvDR08].

The items based on the Game Engagement Questionnaire were scored on a 5-point Likert scale with a range from *No* to *Maybe* to *Yes*. Inspired by the work of Brockmyer et al. the participant's perceived time spent playing were compared to the actual time spent playing. The greater the difference between the perceived time and the actual time playing the game (time accuracy) and the direction of the difference (time estimation) were an indication on how much the participant lost track of time and became engaged. Chaston and Kingstone mentioned that the more a person committed attention to a task the less attention was given to the passage of time from a person's internal clock, which caused the difference between actual and perceived time [CK04].

Satisfaction

The next dependent variable was *satisfaction*. The items of this construct had been inspired by the work of Manresa-Yee et. al and the work of Taylor and Agamanolis. In [MYPVP10] Manresa-Yee et al. primarily looked for satisfaction at how useful their product was for their target demographic. Taylor and Agamanolis held a survey to measure the user experience for a telecare product in [TA10]. Although their focus was on the behavior and reasoning for wearing the product, they got a lot of feedback focussed on the satisfaction of specific aspects.

There were four aspects of the pose game noticeable through the feedback. The first three were the three elements of the feedback, the color changing background, the image of the player and the hint about the desired pose. The fourth aspect was how well the game accepts the player's input pose as the desired pose, by being too lenient, correct or too strict. In addition, would the player expect to enjoy playing the game. Satisfaction was measured with these five items in a questionnaire that focussed on the contribution of each aspect of the feedback to the user experience [TA10]. These items were scored on a 5-point Likert scale with opposing adjectives.

User experience

A general view of the participant's user experience was measured with a qualitative survey during the oral interview, which was inspired by the work of Manresa-Yee et al. [MYPVP10]. The participants were asked to rank the feedback options from best to worst and elaborate on their choices. In addition, the participants were asked what information they thought was lacking in the feedback. The specific questions of the questionnaire and the interview are provided in appendix A.

4.2.3 Dependent variables of user performance

Next to user experience, user performance was also a dependent variable in this study. However user performance is objective in nature and is measured through data logged by the computer. The ISO 9241 standard states that user performance is measured by effectiveness and efficiency [ISO98].

Effectiveness

According to the ISO 9241 definition: "Measures of effectiveness relate the goals or subgoals of the user to the accuracy and completeness with which these goals can be achieved" [ISO98]. For this study it meant that the completeness was determined by the number of participants that completed the tasks[1] for each feedback. The pose game logged whether a participant completed a task with a feedback option.

The accuracy in this study was how close the distance between the participant's pose and the desired pose became. To measure this the feedback data (see Appendix C) of each frame was recorded. The distance between poses in the feedback data was based on a weighted point-to-point distance between the joints of the captured skeleton and target skeleton. The primary joint in the pose had a weight of one (1.0,1.0), while irrelevant joints had a weight of zero (0.0,0.0). Tasks that were completed had a distance of zero, while uncompleted tasks had a distance larger than zero.

Efficiency

The ISO 9241 definition of efficiency is: "Measures of efficiency relate the level of effectiveness achieved to the expenditure of resources." [ISO98]. The physical effort was mostly dependent on the pose that was executed and not on the feedback option. Besides it was not expected that the participants had to put much physical effort in the posing. The only relevant resource in this study was time; the time to either complete or give up on a task.

The pose game recorded the time spent on each task in milliseconds for the most accuracy. The less time necessary to complete a task meant a more efficient feedback. The less time spent on an uncompleted task was also more efficient. As it did not spent time on a task that the participant considered he would never complete.

4.3 Experiment setup

The laboratory setup was simple and shown in figure 4.2. The graphics of the game were projected on a large display. Below this display the Kinect sensor stood elevated from the floor. The participant had enough space so the entire body was visible in the sensor. In addition, when the participant moved he did not bump into anything or anyone.

The experimenter sat at the side behind a desk and observed the participants. At the desk was also the computer that ran the pose game. The experimenter sat outside the view of Kinect sensor. He could control the flow of the experiment when required

[1] A single task meant that a participants assumed a single pose with one of the feedback options. In the experiment the participants did 4 tasks. Each task was defined by the combination of a feedback option and a pose.

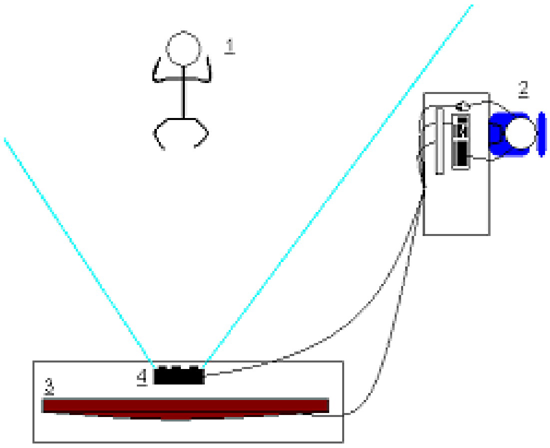


Figure 4.2: This was the layout of the experiment. **1** was the space the participant had to pose in, which was determined by the sensor's field of view. **2** was the experimenter's position. **3** was the large display that displayed the game. **4** was the Kinect sensor.

by using the keyboard. For example, when the participants gave up on completing the task.

4.3.1 The game setup

The poses of the participant needed to be recognized as input. This was done by means of computer vision using the Xbox360® Kinect™ sensor from Microsoft and the OpenNI framework[2]. The sensor, OpenNI framework and the game's preprocessor and data processor (see Appendix C) captured and transformed the captured pose to feedback data that could be used by the game model.

Feedback data for the game model was created by comparing two pose descriptors (a previously recorded descriptor and the descriptor of the captured pose). The model adjusted its feedback settings and the feedback was provided to the participant. Then the game checked if a match was found. In case that a match was found a reward image was created with the participant in it and the next desired pose was selected. If there was no new desired pose to be selected the game was completed and the game continued to showing the reward images.

4.4 Experiment procedure

Before the experiment the participant was welcomed. The experimenter explained to him the purpose of the study and what data was going to be recorded. The experimenter told him about the general flow of the experiment and that the participant had to move to assume the correct pose desired by the game. The game would hint on the desired pose and how far the participant was from the desired pose through the feedback. The experimenter did not tell him what the exact feedback was in order to maintain a certain element of surprise, which could lead to pleasure [BH05].

The experimenter told the participant that he would have to complete four tasks with four different feedback options. Whenever the participant wished to give up on completing the task and continue to the next he needed to notify the experimenter. The

[2]see `http://www.openni.org`, last viewed 2-11-2011.

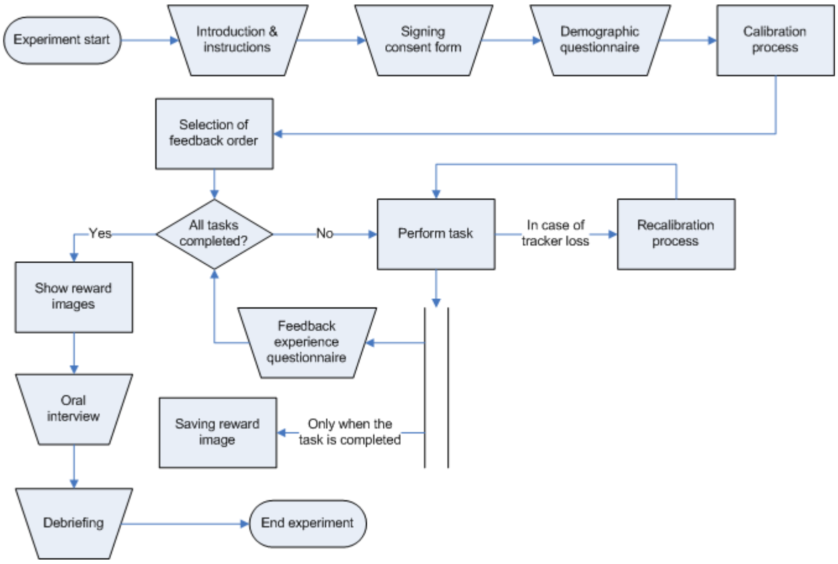


Figure 4.3: This flowchart shows the flow of an experiment. This flowchart assumes that the participant continues playing until all the tasks are completed or given up. If the participant gives up on the experiment all his data is discarded and the experiment terminates immediately.

experimenter also told the participant that he could quit participating in the study at any time.

The participant signed a consent form that any data gathered during the experiment could be used for this study. The data would be reported on, but was completely anonymous and untraceable to the participant. Next the participant filled out a written demographic questionnaire and the experimenter assigned him to a group.

The experimenter started the experiment with the feedback order respective to the participant's group. The participant tried to assume the correct pose based on the information provided by the feedback. Once he had completed or given up on the task he had to fill out a written questionnaire pertaining to the feedback option. The experimenter would tell the game to start the next task when the participant was ready to continue. This process was repeated three more times for each feedback option.

The participant could also choose to give up on completing the task. He then notified the experimenter, who asked him first if he was sure. When the participant was sure that he wanted to give up the experimenter told the game that the participant gave up. Just as completing a task the experimenter presented the participant with the questionnaire and told the game to continue to the next task, when the participant was ready.

When all the tasks were done and the respective questionnaires were filled out the participant might view his reward pictures. After this, the experimenter would take an oral interview asking the participant to compare and rank the various feedback options.

Finally the experimenter debriefed the participant and thanked him for his participation. The experimenter made a copy of the data files containing the logs of the pose game. The entire experiment would be about a half hour in length.

4.5 The subjects

The subject pool was from the University of Twente. They were either lecturers, research staff or students at the university. The age ranged from 18 to 65 years old. There were both females and males participating, although the number of males participating was expected to be larger than the number of females.

The subjects' heights were ranging from 155 cm to 200 cm. Some of them had experience using the Microsoft Kinect sensor, while most would not have experience with it.

The participants were divided at random over the groups. They were told that they needed to move their limbs and torso; and that the position in the room mattered. The participants did not know what feedback to expect.

4.6 The hypotheses

Based on the research questions and the methodology to measure the constructs, these hypotheses were defined.

Hypothesis 1. The feedback options score differently on pleasure.

The first six items of the questionnaire provided the score on pleasure for the respective feedback option. These items were based on the Semantic Differential Scale, which was a validated scale [BL94] and validated to measure the same.

The hypothesis could be accepted when a statistical analysis of pleasure showed a significant effect for feedback.

Hypothesis 2. The feedback options score differently on engagement.

Adapted questions from Brockmyer's Game Engagement Questionnaire provided the score for engagement. The scale was validated and Brockmyer et al. stated that the scale measured four levels of engagement [BFC+09]. These were in order from low to high engagement: immersion, presence, flow and absorption. Immersion and presence were scored with one question each. Flow and absorption were scored by three questions for the respective category.

When there was a significant effect found for feedback with a statistical analysis of the score over all the items of engagement this hypothesis could be accepted.

Hypothesis 3. The feedback options score differently on satisfaction.

The final five items of the questionnaire measured the participant's attitude towards (several aspects of) the pose game [ISO98]. Three items determined whether the participants were positive or negative towards their respective aspects of the feedback. One item determined if the participants considered the game too difficult or easy. And the last item determined whether the participants enjoyed the pose game.

The hypothesis could be accepted when the score for satisfaction provided by all the items showed a significant effect for feedback in a statistical analysis.

Hypothesis 4. The feedback options measure differently on effectiveness.

The purpose of the pose game was to get a participant to assume a pose. The effectiveness was thus how well the feedback option could fulfill the task; and how close the participant came when he failed [ISO98]. The task completion took all the cases into account, while the accuracy would only consider the cases for the feedback where the participant failed to complete the task.

It was expected that one or more feedback options would have too few participants giving up on the task for a reliable measure of the accuracy. Thus only the task completion would be considered for the hypothesis. When a statistical analysis showed a significant effect for feedback on task completion this hypothesis could be accepted.

Hypothesis 5. The feedback options measure differently on efficiency.

The efficiency was determined by the means of the actual time played for all the tasks that were completed; and the means of the actual time played of all the tasks that had been given up for each feedback.

Similar to effectiveness it was expected that one or more feedback options would have too few participants giving up on the task for those to be considered. Thus efficiency only considered the mean times of the tasks that were completed. The hypothesis could be accepted when significant differences were reported between the mean times to complete a task for the feedback options.

Chapter 5

Results

5.1 The participants

A total of 28 participants participated in this study. The participants were divided at random over the groups with 7 participants in each. All the participants were recruited from the University of Twente.

Five participants (18%) were female and 23 participants (82%) were male. Ten participants (36%, one female and nine male) had experience in using the Kinect prior to the experiment.

The average age was about 33 years old with a standard deviation of 10,059 years. The youngest participant was 24 years of age and the oldest was 59 years. The average height of the participants was about 181 centimeters with a range of 155 centimeters for the smallest participant and 198 centimeters for the tallest. Table 5.1 shows the demography for each group.

	#Gender		#Kinect used		Age (in yrs)		Height (in cm)	
	Female	Male	No	Yes	Mean	Std.Dev.	Mean	Std.Dev.
Group 1	2	5	5	2	35,86	13,284	180,29	9,621
Group 2	0	7	4	3	30,71	8,015	186,14	7,777
Group 3	2	5	4	3	28,00	2,000	174,29	14,930
Group 4	1	6	5	2	37,14	13,692	184,29	8,261
Total	5	23	18	10	32,93	10,059	181,25	10,960

Table 5.1: The demographic values for each group.

5.2 Statistical analyses

The analyses focused on the differences in scores between the feedback options. The data had been analyzed by using a repeated measures ANOVA for within-subjects experiments. The combined statistical results can be found in Appendix B.

During the experiment the order was varied to balance the results in four different groups. The group order was an independent between-subjects factor. The four feedback options of the pose game were an independent within-subjects factor. The participants were the second within subjects factor dependent on the group order factor. The degrees of freedom for feedback was three and the interaction of feedback with participants was

72. Whenever Mauchly's test of sphericity showed significant violation of sphericity, the degrees of freedom were corrected with a Huynh-Feldt correction.

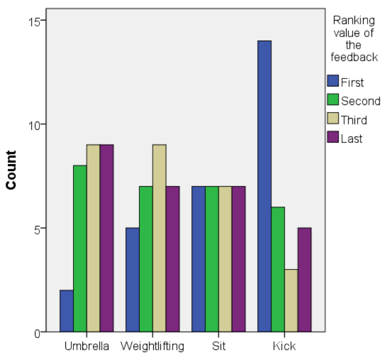
However some variables, such as the efficiency with user performance, had an additional constraint that the task needed to be completed or not. In these cases the sample sizes for each feedback option was not guaranteed to be equal (smallest being $N = 2$) over the feedback options. A one-way ANOVA was used to analyze these variables. The feedback options were in this case a between-subjects factor and the completed or uncompleted tasks were the within-subjects factor. A Levene's statistic was used to determine the homogeneity of the variance.

In the both statistical methods the results were considered significant when $p < 0.05$ and marginally significant when $p < 0.1$.

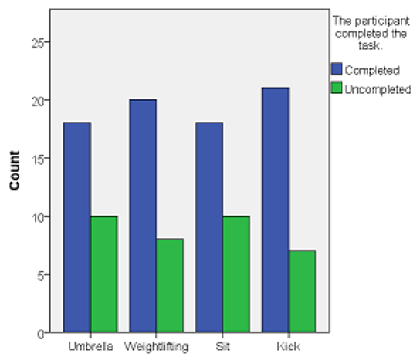
5.3 Poses

The poses were not considered a factor in the statistical analyses due to the fixed order in which they were presented to the participants. However this section will give a quick overview of cross reference analyses of the poses.

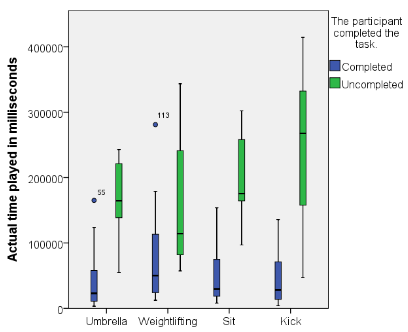
The results in figure 5.1b show that each pose was just as difficult as the others. There was not much difference in the number of completed and uncompleted tasks between the poses. This shows that none of the poses was easier or more difficult to do than another. In the actual playing time there was also not much difference between



(a) Rankings



(b) Completion



(c) Playing time

Figure 5.1: These charts show the results cross referenced with the poses: **(a)** Is a cross reference with the rankings. **(b)** A cross reference with the task completion. **(c)** The mean time played with each pose in case of completion and failure.

the poses (see figure 5.1c). Participants seemed on average to take the longest time in completing the weightlifting pose, while this pose seemed to be the one that participants gave up fastest on. When participants were unsuccessful they tried the longest on the kick pose. However in all these cases there was a large variance to consider.

The more remarkable results can be found in figure 5.1a. It seems as if the umbrella pose was favored the least by the participants and the kick pose was favored the most. The sit pose gave an equal score over all the rankings. And the weightlifting pose seemed to receive the most *third* ranks. However it should be considered that the order of poses was never varied over the participants. For every participant, the umbrella pose was the first pose they needed to assume. The weightlifting and sit poses were respectively the second and third task to be performed. While the kick pose was always the last task. The rankings could also be the result of order in that the participants were becoming more familiar with the experiment. In addition the rankings were asked of the participants after the experiment and the last task was the freshest in their minds.

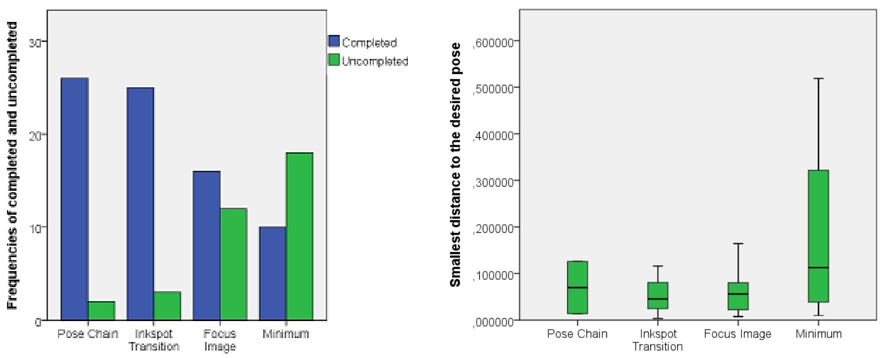
5.4 User performance

The second research question asked which of the four feedback options had the best user performance. The performance was determined by the effectiveness and efficiency. The effectiveness was determined by the number of tasks completed and how close the distance between the poses was with the uncompleted tasks. The efficiency was determined by the time the subjects used to complete or give up on the task [ISO98].

No pose was too difficult to perform with any of the feedback options. The results show that within each group there was at least one participant who completed all four of the tasks that the experiment presented them. Each participant completed at least one of the tasks.

5.4.1 Effectiveness

According to hypothesis 4 there were differences reported in the effectiveness of the feedback options. This hypotheses was tested with a repeated measures ANOVA (see Section 5.2) with the *task completion* as the dependent variable. Task completion could



(a) Task completion (b) Pose Accuracy

Figure 5.2: These charts show the effectiveness for each feedback option: **(a)** The count of tasks completed and uncompleted. **(b)** The accuracy of uncompleted tasks in nearing the pose.

only have the values *completed* and *uncompleted*. The analysis showed that the hypothesis was accepted due to significant within-subject effects between the feedback options on task completion ($F(3.000, 72.000) = 14.169$, $p < 0.001$ with a Huynh-Feldt correction ($\varepsilon = 1.000$)).

Both the pose chain and inkspot transition feedback had significantly more tasks completed than both the focus image and minimum feedback (see Figure 5.2a and Table B.2). The participants also commented that the pose chain feedback was the clearest to understand and the easiest to perform. The inkspot transition feedback option had an abstract shape that turned into a silhouette at its last stage. Once in the final stage the inkspot transition feedback was the same as the pose chain feedback, which explained the similarity in task completion.

Interestingly, was the minimum feedback option which was the only option with more uncompleted tasks (18) than completed (10) tasks. However it was also the only option that did not give a hint to the expected pose. The participants commented that this was frustrating. Only a few appreciated the difficult level of challenge.

The *accuracy of nearing the pose*[1] was based upon those samples that had not completed the pose. Thus it had been analyzed using a one-way ANOVA with the real valued accuracy as the dependent variable. Levene's statistic showed that equal variance could not be assumed ($p < 0.05$) and the ANOVA test showed no significant differences in the mean between the feedback options ($F(3, 31) = 1.750$, p = n.s.).

However the minimum feedback option was the least accurate in nearing the pose with the largest mean distance accuracy. Its standard deviation ($\sigma = 0.30490$) was also larger than the standard deviation of the other feedback options (see Table B.3). The experimenter observed that some participants were reasonably close to the desired pose, while others hardly got the background to change color.

The inkspot transition feedback had the smallest mean accuracy, but the difference with the pose chain and focus image feedback was very small. The final stage of the inkspot was the same as the pose chain feedback. However the experimenter observed with the pose chain feedback that participants were keeping the pose two dimensional for the arms. Posing two dimensionally meant that the participant moved the arms up and down to the side of the body instead of also forward and backward. The focus image feedback seemed to be the most consistent with the smallest standard deviation ($\sigma = 0.04690$, see Table B.3). However it should be noted that the pose chain and inkspot transition feedback had a small sample size of respectively $N = 2$ and $N = 3$ (see Table B.1).

5.4.2 Efficiency

To analyze the efficiency of the feedback options a one-way ANOVA (see Section 5.2) was used, because the differences between the feedback options in the time the participants played the game had to be separated in completed and uncompleted categories. The actual time played was used as the dependent variable for both categories. These results can be found in Tables B.5 and B.7 of Appendix B.

The participants commented that the pose chain feedback was the clearest and easiest. This was reflected in the task completion and efficiency of this feedback. The mean time spent playing when a task was completed was the shortest for the pose chain feedback ($\mu = 41.20$, $\sigma = 34.544$). However hypothesis 5 states that differences between the

[1] The accuracy of nearing the pose was based on recorded feedback data calculated for the inkspot (see Appendix C Equation C.4).

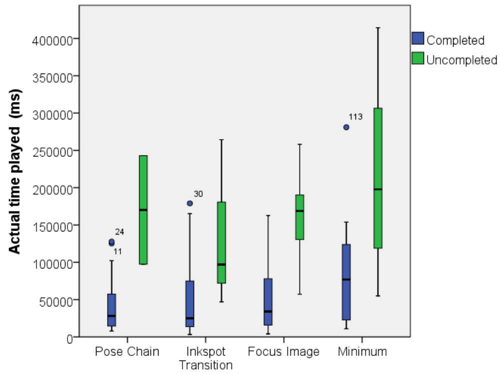


Figure 5.3: This chart shows the efficiency for each feedback option. The time it took for the participant to complete the task or give up.

feedback options were expected. Levene's statistic showed no homogeneity in the variance ($p < 0.05$) and there was only a marginally significant difference reported between the feedback options for the actual playing time of completed tasks ($F(3, 73) = 2.525$, $p < 0.1$). Thus hypothesis 5 cold not be accepted and was rejected.

The pose chain feedback required the participant to pose in three poses, but this had little influence on the completion time. The experimenter observed the participants spending most of the time finding the correct position for posing. Once this position was found the other poses in the chain were assumed relatively faster. For the inkspot transition feedback the challenge was 'to find the correct direction' to pose in, as the participants called it. And with the focus image the participants mentioned that it required some creative thinking, because the feedback presented you with an object. The participants did comment that they were confused at first. With the inkspot transition and focus image feedback options they first spent some time considering the hint that they were given. With the minimum feedback option the participants were confused as well, but most of them first waited to see if they would not get a hint after all.

The minimum feedback was the most inefficient option with $\mu = 92.94$ and $\sigma = 82.464$. The participants commented that this feedback was too difficult and more or less they had an infinite possible directions to start posing in, without a hint on where to start. The feedback relied a lot on 'lucky shot' tries to obtain the pose. This was somewhat confirmed by the fact that the fastest completion time on record for this feedback was about 10 seconds. The respective participant commented that he was surprised and had no idea what he had been doing to succeed.

For the uncompleted tasks it was the inkspot transition feedback that took the least amount of time ($\mu = 136.06$ seconds). Thus the participants gave up the fastest with this option. With the pose chain ($\mu = 170.17$ seconds) the participants tried on average more than 30 seconds longer than the inkspot transition feedback. However even with Levene's statistic showing that equal variance could be assumed (p = n.s.), the one-way ANOVA results for the efficiency of uncompleted tasks showed no significant differences ($F(3, 31) = 1.331$, p = n.s.).

The minimum feedback did not give a hint to the pose and yet the participants still tried on average for more than three and a half minutes to find the pose. This was longer than any of the other feedback options. Thus the minimum feedback option was also very inefficient when the player could not complete the task. Participants spent the most time before giving up on the game with this option.

Research question 2 asked which feedback option had the best user performance (see Section 4.1). The answer to subquestion 2A, “which feedback is the most effective”, were the pose chain and inkspot transition feedback. In task completion the pose chain feedback option had one more completed task than the inkspot transition feedback, which was not a significant difference. Yet the accuracy of pose distance with the uncompleted tasks the inkspot transition feedback seemed to do better than the pose chain, even though the difference was not significant. However significant within-subjects effects were reported on task completion thus hypothesis 4 was accepted.

No significant differences had been reported for the efficiency of completing tasks, which meant that hypothesis 5 was rejected. Although the pose chain was the most efficient when it came to completing a task. And this answers research question 2B of which feedback option is the most efficient. Although the inkspot transition feedback seemed to take the least time when being unsuccessful.

The feedback option with the best user performance was the pose chain feedback. It had the better measured task completion and was measured to be the most efficient in completing the task. The accuracy of pose distance and the time spent with uncompleted tasks was based on too small a sample size for the pose chain and inkspot transition feedback to matter for performance. The minimum feedback was definitely the worst in user performance. It had a low task completion and had the worst pose accuracy. It also cost the most time no matter whether the task was completed or not.

5.5 User experience

“Which of the feedback options provides the best user experience” was asked by the first research question. Because user experience is too broad a concept to be measured [Has05], this study focused on the constructs pleasure [BL94], engagement [BFC$^+$09], and satisfaction [MYPVP10, ISO98].

In the interview the participants were asked to rank the feedback options according to their experience. The participants often commented that the minimum feedback was frustrating and way too difficult, which was clearly visible in the results. The minimum feedback was significantly ranked the worst feedback by 20 of the 28 participants (see Figure 5.4 and Table B.10 in Appendix B). The pose chain had the most votes ranking it as the favorite option. The reason the participants gave, is the pose chain being the

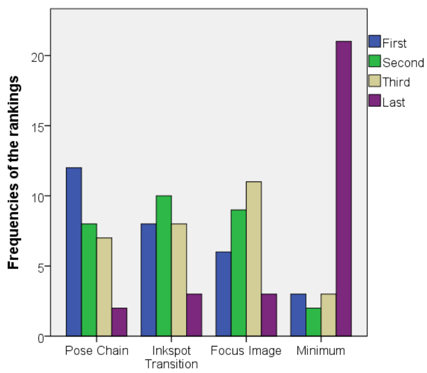


Figure 5.4: This chart shows the frequency of how often each feedback option was given a rank by a participant. The pose chain feedback seemed to be the favorite, followed by the inkspot transition feedback as second and the focus image as third. The minimum feedback was significantly ranked the worst.

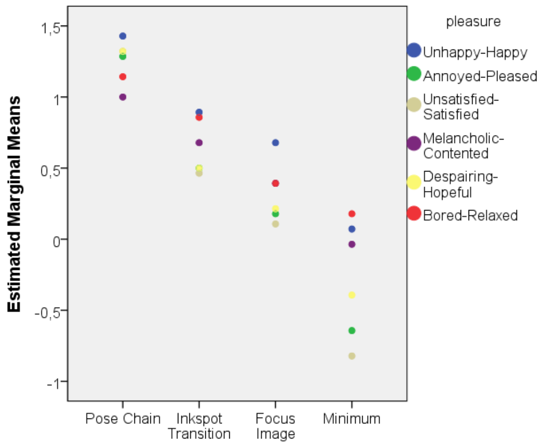


Figure 5.5: The estimated means for each item on pleasure with each feedback option.

most clear in what had to be done.

5.5.1 Pleasure

For each feedback option pleasure was scored by the first six items of the questionnaire. The range of score went from -4 for one end of the adjective pair to +4 for the other end, due to the adjectives being opposites from each other. In a repeated measures ANOVA (see Section 5.2) these six items provided the score on the dependent variable *pleasure*. According to hypothesis 1 the feedback options score differently on pleasure. This hypothesis was accepted as there were significant within-subjects effects reported between the feedback options on pleasure ($F(3, 72) = 8.912$, $p < 0.001$). The pose chain feedback scored significantly higher than the three other options (see Table B.12 in Appendix B).

There was a definite order in scoring recognizable from figure 5.5, which provided an answer to research question 1A (see Section 4.1). The minimum feedback was the worst of them and even had a negative score of $\mu = -0.274$. This feedback was thus on average associated with a negative emotional state. Which was expected considering the negative comments most of the participants made about this feedback.

The inkspot transition ($\mu = 0.649$) and focus image feedback ($\mu = 0.327$) were respectively second and third in the scoring. This was interesting because the participants that favored these feedback options gave as reason that they were more challenging or required more creativity than the pose chain option. The pose chain was considered by them to be too easy. Blythe and Hassenzahl state that the most pleasurable experience is the one with the best flow [BH05]. And flow is the balance between challenge and skill [MC96]. The inkspot and focus image provided a closer balance than the pose chain option. Although the focus image option was commented by some to be too difficult. However they did not provide the most pleasure.

In this study it was the feedback option that was the clearest to understand which provided the most pleasure. In answer to research question 1A, the pose chain option scored significantly better than the others ($\mu = 1.250$, $p < 0.05$). Mainly because it was clear on what needed to be done and did not require "additional onscreen instructions" as some participants mentioned. Some commented that the other feedback options tended to create some confusion on what to do at first.

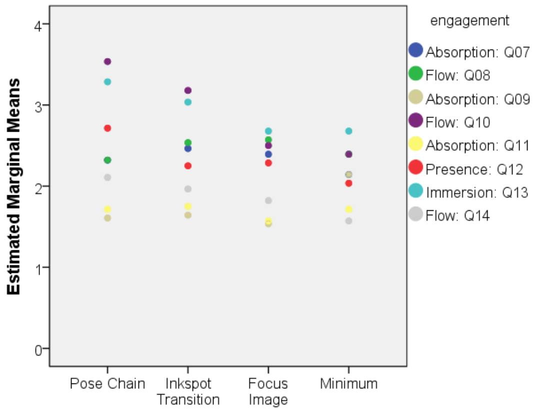


Figure 5.6: The estimated means for each item on engagement with each feedback option. Each item is indicated by the respective category and refers to the questions in the questionnaire (see Appendix A)

5.5.2 Engagement

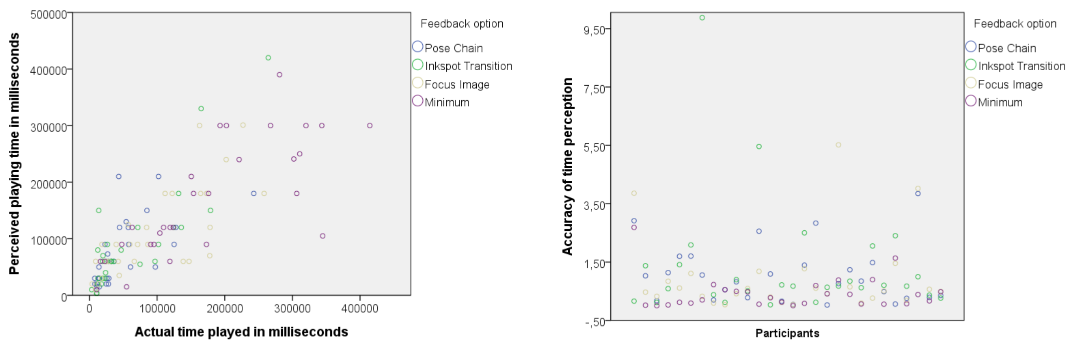
Engagement was measured with two parts. The first and main part was a questionnaire scoring four categories of engagement [BFC^{+}09]. For two categories there was only one item providing the score. The other two categories each had three items that provided the score for their respective categories as dependent variables in a repeated measures ANOVA (see Section 5.2). The range of the score went from one to five. The results are presented in Appendix B.

The second part served as a backup to answer research question 1B and was the distortion in time perception. The participants were asked to note down the number of minutes and seconds they thought they had played the game. The over-/underestimation and the accuracy of their perception to the actual playing time was calculated. No distinction between completed and uncompleted tasks was made because the participants either succeeded or chose to give up by themselves.

Hypothesis 2 states that differences in results between the feedback options were expected. To determine this hypothesis' validity the score of engagement was analyzed which was provided by all the items of each category. This hypothesis was accepted due to the significant reported within-subjects effects between the feedback options on engagement ($F(3, 72) = 3.994$, $p < 0.05$).

The highest category, absorption, showed no significant differences between the feedback options (see Table B.20). In the other three categories some differences could be found. The pose chain feedback seemed to score better than the other options (immersion $\mu = 3.286$, presence $\mu = 2.714$ and flow $\mu = 2.655$). In the category presence it scored even significantly better than the three other options, while for flow and immersion it scored significantly better than the focus image and minimum feedback options. The inkspot transition scored second on immersion and flow (respectively $\mu = 3.036$ and $\mu = 2.560$), while the focus image feedback scored second on presence ($\mu = 2.286$).

To determine the time distortion the values for time estimation and time accuracy had been calculated. These values were analyzed using a repeated measures ANOVA (see Section 5.2). The means of the feedback options showed that all the options were on average overestimated, thus on average the participants considered that they had played longer than they actually did (see Table B.21). The analysis showed marginally significant within-subjects effects between the feedback options on the time estimation



(a) Perception versus actual

(b) Time accuracy

Figure 5.7: These charts show the distortion of time measured with the pose game. **(a)** is the perceived time set out against the actual playing time for each feedback option. **(b)** is the accuracy of time perception for each participant with each feedback option.

($F(1.984, 47.623) = 2.948$, $p < 0.1$ with a Huynh-Feldt correction ($\varepsilon = 0.661$)). Though the minimum feedback was significantly smaller in its overestimation than the other three feedback options (see Table B.22). The number of participants that overestimated to the number that underestimated their playing time with the minimum feedback was almost even, while the other three feedback options were mostly overestimated.

In the accuracy of time perception there were some marginally significant within-subjects effects reported between the feedback options ($F(1.907, 45.759) = 2.504$, $p < 0.1$ with a Huynh-Feldt correction ($\varepsilon = 0.636$)). The participants were significantly more accurate with the minimum feedback option ($\mu = 0.433$, $p < 0.05$) than the other options (see Table B.24). The least accurate was the inkspot transition feedback ($\mu = 1.331$), due to some very large outliers as is shown in figure 5.7.

With the minimum feedback options the participants seemed to be able to spend the most attention to a correct duration judgment and the least attention with the inkspot transition feedback which caused the larger time disruption [BZ97]. The participants focussed the most attention on the game with the inkspot transition feedback.

Research question 1B asks, which feedback option got the player the most engaged in the game. There was not any apparent order of the feedback option in the scoring. The pose chain feedback seemed to score the best. However it must be noted that the items for flow and absorption should correlate on their respective categories. But there no longer seemed to be a correlation between the items of the same category. And looking at the distorted time perception the inkspot transition feedback seemed to get the most attention of the participants and thus the most engagement.

On the subject of engagement there were no definite answers. The results seemed to be influenced by other factors of the experiment, which made it difficult to accurately measure engagement. This will be discussed further in Subsection 6.2.1.

5.5.3 Satisfaction

Satisfaction had been measured with five items each highlighting an aspect of the feedback and the pose game. Each item ranged from -2 to +2 in the score on a five point Likert scale, due to opposing terms. Each item had been separately analyzed using a repeated measures ANOVA (see Section 5.2).

Hypothesis 3 specifies that differences were expected between the feedback options on satisfaction. This could be accepted when the score on satisfaction showed significant

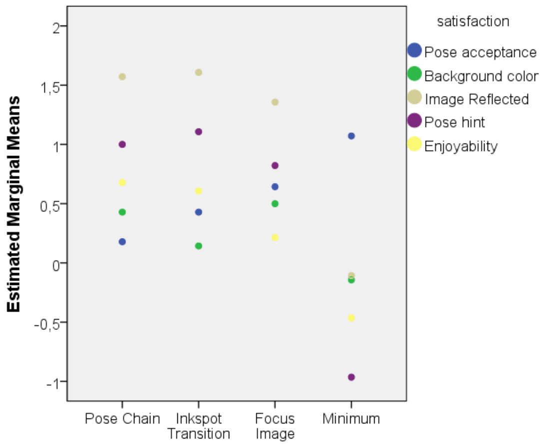


Figure 5.8: The estimated means for each item on satisfaction with each feedback option.

within-subjects effects. The five items of satisfaction provided the score for the dependent variable satisfaction. In the score on satisfaction significant within-subjects effects were reported ($F(2.586, 62.065) = 18.404$, $p < 0.001$ with a Huynh-Feldt correction ($\varepsilon = 0.862$)). Thus the hypothesis was accepted.

The minimum feedback option scored in three different items significantly worse than the other options and in a fourth item the pose chain was considered significantly better than the focus image and minimum feedback options. The background aspect of the feedback was the only aspect without significant differences between the feedback options. The background scored the highest for the focus image ($\mu = 0.500$), while the score for the minimum feedback option ($\mu = -0.143$) was the lowest. It seems the participants considered the color changing background the most useful with the focus image feedback. However the differences were small and not significant, so that the usefulness of the color changing background could be considered similar between the feedback options (see Subsection 6.2.3).

A negative score on the pose acceptance for the feedback meant that the pose was accepted before the participant considered himself to be in the right pose; being lenient. While a positive score indicated that the participant thought he was in the right pose but the game did not accept it as the right pose; being strict. On average none of the feedback options were too lenient when accepting the pose.

The minimum feedback ($\mu = 1.071$) was considered to be the most strict (see Table B.25). Some participants got the background changed to a green color and yet it was not green enough. Others did not even get close to the green. This frustrated them. Ten participants completed the task and only a few were really fast (< 30 seconds).

Remarkably the pose chain feedback ($\mu = 0.179$) seemed to have the best pose acceptance. During the experiment the participants commented that they had copied the pose, or that they were completely enveloped by the silhouette and still the pose was not accepted. With the inkspot transition feedback ($\mu = 0.429$) these comments were also made, but not as often. The two dimensional posing with the pose chain feedback was an obstacle especially for the first two poses in the chain.

For all the other aspects the minimum feedback seemed to have scored significantly worse (see Table B.25). This was not surprising considering that most participants gave negative comments about it and were frustrated with this option. Thus the scoring of their attitude towards the various parts reflected this.

The inkspot transition feedback ($\mu = 1.107$) seemed to provide the most positive experience for the pose hint. Closely followed by the pose chain feedback ($\mu = 1.000$). The focus image feedback ($\mu = 0.821$) was a close third in the score. However there was no significant difference between the three (see Table B.25).

Seeing your image on screen showed similar results, although on average a bit higher score. This suggested that the reflection was only as useful as the hint to the pose that they had to work with. However there was some usefulness in the fact that the game did recognize them from the camera image and could compare it to the desired pose.

On overall enjoyability the pose chain feedback ($\mu = 0.679$) scored the best, closely followed by the inkspot transition feedback ($\mu = 0.607$). The focus image feedback ($\mu = 0.214$) scored marginally significantly worse (see Table B.26). Giving credence that the feedback was still doable, but was often too big a challenge with the parts of the feedback presented in this experiment.

To answer research question 1C: "With which of the feedback options are subjects most satisfied?". This was both the pose chain and the inkspot transition feedback. The differences between these options over all the items was very small and not significant, so that they could be considered similar. Sometimes the pose chain scored a bit better and on others the inkspot scored a bit better. The minimum feedback provided by far the least satisfaction, as it scored significantly worse on four of the five items. The focus image feedback scored similarly to the pose chain and inkspot transition feedback on some of the items. However there was a significant difference with the pose chain feedback option on the pose acceptance.

To summarize, the hypotheses 1 and 3 were accepted as differences were reported. In general the minimum feedback was the most different from the others and scored often worse. There were differences found between the feedback options for engagement and thus hypothesis 2 was accepted. However there were some other factors that prevented an accurate measurement of the engagement. And items that should correlate on the same category, no longer do so.

Research question 1 asked which feedback option provided the best user experience. Due to the lack of accurate measurement for engagement the answer was based on only pleasure and satisfaction. The minimum feedback was definitely given the worst user experience. The participants were generally negative about it. And only a few actually favored this option. But most were too frustrated with the lack of information given by this feedback.

For satisfaction both the pose chain and inkspot transition feedback were similar. However for pleasure the pose chain feedback option was significantly better than the other options. The participants also ranked the pose chain most often the best, while the inkspot transition feedback was most often ranked second. Thus the pose chain feedback provided the best user experience. This was remarkable because the participants commented that the inkspot transition feedback was the more challenging of the two. However the pose chain was the clearer one and the participants immediately knew what to do.

5.6 Observations and comments

One of the participants commented that he recognized the game show 'Hole in the Wall' from television in this pose game, especially with the pose chain feedback option. Analyzing more of the participants' comments showed defining characteristics for each feedback option. The pose chain feedback was clear and easy. The inkspot transition was challenging, but doable. The focus image feedback required the participant to be creative. And the minimum feedback was for most participants frustrating and too difficult. Although a few participants considered the pose chain feedback to be too difficult or the minimum feedback too easy.

In general the better a participant could accomplish a task the higher the ranking was. However one participant that failed the inkspot transition feedback and one that failed the minimum feedback ranked both feedback as their favorite, arguing that it was the challenge what they liked.

Most participants were very negative about the minimum feedback. Some called it "ridiculous". Most participants first looked confused and just stood still in front of the camera hoping there would be another hint. With the inkspot transition and the focus image feedback the participants were confused at first as well, but they started to move to discover what the hint meant and the feedback had to offer.

The minimum feedback did not provide a hint, which made it so difficult. Participants commented that finding the right pose was a matter of getting a 'lucky shot'. Only the color changing background offered some information to the pose. The colors were well understood by the participants as red meant wrong and green was correct. However they did suggest that a progress bar or indicator would be better, because color-blind people have trouble distinguishing the red and green colors. In addition a progress bar would provide a clear ending and distance to how far the pose was still removed from the desired pose. With all the feedback options the participants commented that they did not know "how green, the green has to be" and they would have liked to see an example.

Most participants commented that they often found themselves in a position where they were stuck and had a hard time figuring out how to improve their pose. The participants were close with posing and only had a single joint too far away from its target position. Interestingly enough they did not want more information about the pose that was the goal, but information about the body parts that were still wrong. "What needs to be improved." The participants made this comment for all of the options. Even the minimum feedback option, although a few would have liked to have a hint to the pose as well.

Most participants did not recognize the three poses of the pose chain as being sequential poses from a single movement. After they had successfully assumed one pose in the chain, the participants returned to a neutral pose in order to assume the next pose in the chain. Only one participant had seen the connection between the poses in the chain. A few participants considered the inkspot transition feedback to be the same as the pose chain feedback, only with a single pose. After the experimenter explained the feedback options they understood the difference. However they did comment that the transitions between the inkspots went so fast, that they did not notice it.

The experimenter observed that the use of the leg was not the most obvious action to make. With the exception of the pose chain feedback, it took a relatively long time before the correct leg was raised to assume the kick pose by a lot of the participants. With the minimum feedback some participants gave up before considering raising either leg. While with the inkspot transition feedback the participants were fixed on one leg. With the kick pose the focus image's creativity was very apparent. Some participants violently kicked towards the ball immediately, while others first tried to hold the ball in their hands.

A number of participants answered that this game would be a lot of fun playing with more people, when they were asked what further improvements they would like to see.

Some would also like to have the addition of sound with this game and maybe some sounds to give cues for the right direction of posing. The sound could also be used for a success message when the pose has been found. The success message was something a few participants missed as well.

Chapter 6

Discussion

In this study the user performance and user experience were measured for the four feedback options implemented for the pose game. Some results were expected and some were unexpected. In this chapter these results will be discussed in more detail.

6.1 User performance

That the minimum feedback had the worst performance was not unexpected. It was like being dropped in an unfamiliar environment without a map or a compass. This would be a challenge, but it could also be too difficult. Searching in the dark is a task that cannot be completed fast.

It was not unexpected either that the pose chain should have the best performance. In essence this feedback option showed a silhouette and the player had to copy the pose shown in the silhouette. Remarkable though was the close match with the inkspot transition feedback. In this option the participant had to search for the silhouette to copy. Without a hint in what direction they need to look.

The pose chain feedback required that the participant should assume three poses. The participants did comment that the trick was to find the correct position in the room first and the other poses follow up more easily. Some participants mentioned that they had not seen the transitions of the inkspots because they moved the required limbs (often arms) very fast towards the required pose position. This suggested that the desired pose still did not differ enough from the base stance. Copying a silhouette seemed to be a straightforward and simple game mechanism in the pose chain feedback. However such a simple mechanism also came with assumptions and when incorrect they could provide some difficulty in completing a pose. These assumptions were the spacial dimensions of posing and what was the primary joint in the pose (see Section 4.2.3).

The first assumption was the spacial dimensions of posing. The experimenter observed that with the pose chain feedback participants were posing two dimensionally. This meant that they moved their arms and legs up and down at the side of their body in full view of the camera. However with the first two poses in the chain the limbs pointed towards the camera. This created a difference in the two dimensional body proportions between participant and silhouette on top of the difference between recording and live capture (see Appendix C).

The second assumption was to assume the wrong joint as the primary joint in the pose. The experimenter also observed that participants used their feet as the main joint for the pose. However in a pose there was only one joint in the skeleton that had a

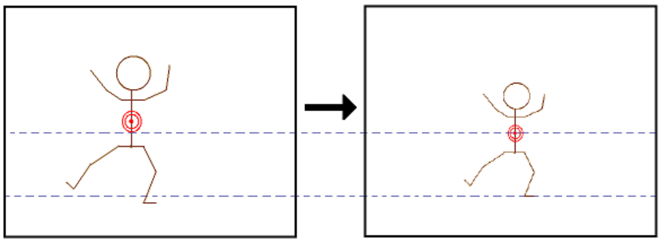


Figure 6.1: This sketch shows the problem with 2D measurements of 3D manipulations. The sensor was at a fixed position. As the person moved away from the camera they became smaller, the feet joints rose, but the center of mass lowered. Thus the joints above the center lowered faster than the lower joints rose.

weight of one (1.0,1.0) for the pose distance calculation (see Appendix C). This was the primary joint and was considered to be the most essential to the context of the pose. In most poses this was not the foot that was standing firmly on the ground.

The focus image and minimum feedback were not troubled by these complications as they had no joint to target or falsely assume as the primary joint. The inkspot transition feedback might eventually also be copying a silhouette, but to get that hint the participants often found the right position and the joint that mattered. In addition these options only considered the final third pose in the chain, which did not point any limb towards the camera.

In addition to the wrong assumptions that participants made, there was a technical difficulty that could cause a drop in performance. This difficulty was caused by the transformation process of sensor to feedback data that did not account for body proportions. This complication is illustrated in figure 6.1 and caused by two dimensional measurements of three dimensional body manipulations. People could move away and towards the camera to make their size similar to the silhouette's. But the camera was in a fixed position and angle, which caused the center of mass of the person to lower or rise. The transformation process was the same for all the feedback options, thus it had little influence on the results.

Still the pure copying mechanism of the pose chain and the association with the TV show "Hole in the Wall" by a number of participants made this option easy to understand and easy to perform. The pose chain feedback did not leave much in the way of ambiguity for what pose was desired by the game, which the participants commented on for the other feedback options.

6.2 User experience

6.2.1 Engagement

Of the results related to user experience, those for engagement were the most noticeable. Questions were adapted from the *Game Engagement Questionnaire* by Brockmyer et al [BFC+09]. However the items within the categories flow and absorption no longer seemed to correlate. For this study, the number of questions was reduced as not to overload the participants with questions and thus influence the results [DFAB03]. This reduction might have violated the validity of the questionnaire.

Brockmyer et al. validated the questionnaire with the videogame S.T.A.L.K.E.R. And that game was a sedentary videogame, unlike the pose game in this study which was an exertion game. Other studies also showed that exertion could have an influence on a person's mental/emotional state [RN08, MBB10, MDHS10]. However due to the exertion some questions might be interpreted differently. For example the question from the game engagement questionnaire "I feel different" was ambiguous [BFC+09]. It could be approached from a mental and emotional aspect, which was more appropriate for a sedentary game, or from a physical aspect, which was more likely with an exertion game. During the pilot experiment the participants did not ask in what sense they should feel, but with what they needed to compare it to. The participants of the study also commented that they did feel the stress on their muscles when posing, especially with the sit and kick poses. It stands to reason that they approached this question from a physical aspect. In hindsight the game engagement questionnaire might not have been the most appropriate questionnaire to measure engagement.

Another reason why a more accurate measurement of engagement might not have been possible is that the task for the participants was too short. The time it took for a participant to finish a task ranged from four seconds to seven minutes. The shorter playing times were because the task had been completed. And the shorter playing times might not have been enough for the participant to become engaged.

Besides the shortness of the task the experiment was also done in a laboratory setting. The participants were recruited from the university and over half of them had never used a Kinect before. The participants motivation was mostly extrinsic by helping out with the research. Intrinsic motivation is more optimal to get into a flow state than extrinsic motivation [PTF+95]. And the participants were told that a task would only end when they either completed the task or chose to give up on completing. They could quit whenever they wanted and it would even have been a valid action to do for this experiment. There was no additional motivation for the participants to finish a task either. Thus the circumstances made it difficult for a participant to become engaged and it was more than likely that participants had not been engaged in the game for any of the feedback options.

The measure of time distortion was still the most telling for engagement. It was expected that the engagement would result in an underestimation of time, with the most engaging feedback option having the largest underestimate. Chaston and Kingston mentioned in [CK04] that if attention was committed to time estimates the "clicks" of a person's internal clock would pass unimpeded. And the more attention was diverted to other tasks the more these clicks would be slowed down or even would fail to pass.

However the results showed that on average the playing time for each feedback option was overestimated. Tse et al. argued that if a person's attention was focused on a task with time duration cues that time might subjectively be expanded and thus overestimated [TIRC04]. The participants were engaged in an exertion game, where they needed to move their body into the correct pose. Movement is always considered with a direction and velocity. Velocity is in common physics knowledge determined as the distance over time. Thus moving one's body is a task that has its own time indication. In addition, the participants needed to figure out what the hint to the pose meant which was an additional task that diverted attention away from time estimation [OECN10].

The minimum feedback option was significantly different than the other three options in both accuracy and estimation. It had the lowest score and was thus the most accurate. With the minimum feedback the participants were able to put more attention to time judgment than with the other feedback and the pose game was thus given less attention. The minimum feedback was the least engaging. The inkspot transition had the largest accuracy and the largest overestimation. Compared to the other options the most attention was spent on playing the game with the inkspot transition feedback option. In time distortion the inkspot transition feedback seemed to be the most engaging.

It should be noted that a number of participants commented that they rounded their time perception to the whole or half minute.

6.2.2 Pleasure

The participants did not spend a long time playing with each feedback option. And yet with significant results the pose chain feedback scored better than the other feedback options. According to the semantic differential scale this meant that the pose chain feedback was the most pleasurable of the options [BL94].

However considering the definitions that Blythe and Hassenzahl offered this was not entirely correct [BH05]. As previously mentioned the pose game did not absorb the participant, it was even possible that there was no engagement. A single task did not take more than a minute to complete. The use of the game did not change over time or between tasks. Neither had the pose game any relevance to the life of the participant. It did nothing to progress the participant's self either. And the pose game was not played to accomplish a relevant task to the participant. The pose game was short, trivial, repetitive and it offered a distraction to the participant.

In other words the pose chain was the most fun to play and not the most pleasurable. The reason that most participants ranked the pose chain feedback option as the most preferable was that it was the clearest to understand of them all. The participants knew immediately what to do and could reap the rewards quickly. This was advantageous for a short distracting game. However if the game was played with prior knowledge of how it would work and what to expect, then the pose chain might not have been the most fun option.

That the pose chain was the more intuitive option was also clear from the results in the rankings. It was the only feedback option to receive the most favorable ranking of participants when they had to assume the umbrella pose. And the umbrella pose was the first that the participants had to do. None of the other options had a similar result.

Some participants commented that the inkspot transition feedback was more challenging than the pose chain option. And its task completion was similar. It was unexpected that it did not give more pleasure then pose chain feedback [MC96]. The pose chain was favored by the participants because it was the clearest on what needed to be done. The participants would have liked to have more instructions about what to expect. However an explicit choice was made not to reveal the full details of the feedback in the pre-experiment instructions. A few participants also commented that the pre-experiment instructions were boring them and that they just wanted to start.

The focus image was also more challenging than the pose chain feedback. In addition it required more creativity than the pose chain feedback. To some participants it was also too difficult. Finding a desired pose with this option might require too much commitment. And putting in more commitment to the game, would have made it less trivial and less of a distraction, thus less fun.

That is why the minimum feedback was considered to be the least favorable by most participants. It was too difficult and to find the desired pose, either required a *lucky shot* or a commitment to systematically try for each possible pose. This was no longer a distraction and a short fun play. Which resulted in most participants giving up on completing the task and thus not receiving a reward. This was not a good first impression and participants judged it more negatively.

6.2.3 Satisfaction

Blythe and Hassenzahl defined *satisfaction* to be the confirmation of the expected [BH05] and the International Organization of Standards defined satisfaction as “the freedom from discomfort and the attitude towards the use of the product” [ISO98]. Thus as long as the feedback did what the participants expected it to do then the participants were satisfied.

The participants were the most satisfied with both the pose chain and the inkspot transition feedback. They scored really similarly in all aspects. They were also the feedback options that had the best task completion. For the participants the elements of the feedback work as they let them complete the task.

Remarkable was that the focus image also scored similarly for the parts of the participant’s image and the hint towards the pose that was given by the feedback. The overall enjoyability of the focus image was also by a margin not significant compared to the pose chain and inkspot transition feedback. And yet on task completion the focus image scored significantly worse than the pose chain and inkspot transition feedback.

The experimenter told the participants before the experiment that they would know what pose was desired based on the feedback. For the pose chain and inkspot transition feedback the pose was obvious as the participants (eventually) needed to copy the pose presented. For the focus image the desired pose was less obvious. But from the presented object in the image the participants could deduce what kind of pose was desired. The feedback did what it needed to do and confirmed what was expected.

However the focus image scored significantly worse than the pose chain and noticeably worse than the inkspot transition when it came to the acceptance of the pose. The participants commented that the most frustrating part of the focus image was that it did not accept the pose when the participant thought he had assumed the desired pose. Although the acceptance of a pose was the same over all the feedback options as explained in appendix C. For the focus image this part did not meet with the expectations causing some discomfort.

The results suggested that satisfaction was a deficiency part of user experience [HDG10]. As long as things were what was expected then the participants were in general not more satisfied with it. However if it did not work as expected and caused some discomfort then various degrees of dissatisfaction could be distinguished. This was most obvious with the minimum feedback. It lacked a hint to the desired pose and on almost all the questions the minimum feedback scored significantly worse. The participants expected to learn what pose was desired from the feedback. And even if the task was completed the participants still did not know what pose had been expected from them. The minimum feedback just did not meet the expectations of the participants.

The only part of the feedback that the minimum option was not significantly worse was the color changing background. The minimum feedback did have the lowest mean score and the focus image had the highest mean score, but the differences were small and not significant. Some participants also commented that they had not noticed the color change indicating their progress until they had been presented with the focus image or minimum feedback options where the background was necessary to find the correct pose. The participants did comment that they preferred a progress bar or indicator to show how close they were to the desired pose. This suggests that the background was just a background to the participants. It did not have a functional meaning.

Chapter 7

Conclusions

This study had aimed to evaluate the user performance and user experience of four different feedback options with the pose game. In all four options the image of the participant was displayed, like in a mirror. In addition the color of the background indicated the distance between the participant's pose and the desired pose. The feedback options differed in the hint that was given about the desired pose. The minimum feedback did not give a hint. The pose chain feedback showed a series of silhouettes. The inkspot transition feedback first showed a shape that later turned into a silhouette. The focus image feedback showed the image of an object.

In the study we evaluated which of these four feedback options gave the best user experience and user performance. The user performance was determined by measuring the constructs effectiveness and efficiency of the participants with the pose game. For the user experience, the study was narrowed to measuring the constructs pleasure, engagement and satisfaction. It was expected that significant within-subjects effects between the feedback options on their respective constructs would be reported.

For the study 28 participants were recruited and divided over four groups. The participants tried each feedback option with a single pose. A task (assuming a pose with a given feedback option) was finished once the participant found the pose or the participant chose to give up. The survey involved questions from a questionnaire after each task and an interview at the end of the experiment. Additionally, the pose game logged the performance of each participant with each task.

7.1 General results

The results showed that the poses implemented in this game could be assumed with each feedback option as in each group there was at least one participant that completed all four the tasks and the number of completed and uncompleted tasks were similar for all of the poses. The time spent playing with each pose was on average also similar. The poses did not seem to influence the results.

The pose chain feedback option was generally considered to be the easiest and was measured to be the quickest to complete. And the clearest to immediately understand what to do. A few participants even associated it with the television show "Hole in the Wall". Most participants did not recognize that the three poses in the chain described a single movement to the desired pose. They saw each pose as a different pose without a connection to the other poses in the chain. A few participants considered the inkspot transition feedback the same as the pose chain, due to the quick transitions of the inkspots. On average the inkspot transition feedback was considered to be more challenging than the pose chain feedback as it left the player guessing, before eventually showing the silhouette of the pose that needed to be copied.

The focus image feedback was also considered to be more challenging than the pose chain feedback. It required more creativity of the player to assume a pose. It was not always clear what pose was desired, even though participants got to see a picture of an object as a hint. An object can be used in several ways that the player had to consider. Even then, there were variations of the same pose. For some participants it was a bit too difficult. The minimum feedback was also too difficult for most participants. The feedback option did not provide a hint to the pose. This frustrated most participants. According to them the only way to find the pose was with a lot of luck. Only a few participants appreciated this challenge, but most of them were generally negative about it.

7.2 Acceptance of the hypotheses

Some differences between the feedback options were expected for each construct of user performance and user experience. For this study we formulated the following hypotheses:

Hypothesis 1. The feedback options score differently on pleasure.
Hypothesis 2. The feedback options score differently on engagement.
Hypothesis 3. The feedback options score differently on satisfaction.
Hypothesis 4. The feedback options measure differently on effectiveness.
Hypothesis 5. The feedback options measure differently on efficiency.

Hypotheses 1, 2 and 3 were the expectations on user experience. Hypothesis 1 could be accepted when the score on pleasure provided by the respective items from the survey showed significant within-subjects effects between the feedback options. Similar expectations were for the hypotheses 2 and 3 but with the score of their respective constructs provided by their respective items from the survey. The repeated measures ANOVA results showed that all three of the hypotheses were accepted due to significant within-subjects effects on pleasure, engagement and satisfaction (see Section 5.5).

Hypotheses 4 and 5 were the expectations on user performance. The expectations were that there would be significant differences between the feedback options in task completion (effectiveness) and the time playing the game for completed tasks (efficiency). Hypothesis 4 was accepted because a repeated measures ANOVA on the task completion showed that there were significant within-subjects effects between the feedback options. The only hypothesis that was rejected was hypothesis 5. The acceptance of the hypothesis was based on the time it took to complete a task. A one-way ANOVA showed that there were no significant difference between the feedback options in the time it took to complete a task (see Section 5.4).

7.3 Answers to the research questions

The study had two main research questions and five subquestion:

1. *Which of the feedback options provides the best user experience?*
   (a) Which of the feedback options provides the most pleasure?
   (b) Which of the feedback options is the most engaging?
   (c) With which of the feedback options are the participants most satisfied?

2. *Which of the feedback options has the best user performance?*
   (a) Which feedback option is the most effective?
   (b) Which feedback option is the most efficient?

7.3.1 User experience

The first subquestion was which feedback option provided the most pleasure. The pose chain feedback option scored significantly better than the other three options. Thus the pose chain feedback option was the definite answer to question 1A. Although considering Blythe and Hassenzahl's distinction of fun and pleasure [BH05] it is better to say that the pose chain feedback was the most fun. The third subquestion asked which provided the most satisfaction. The answer to question 1C was both the pose chain and inkspot transition feedback options. These feedback options scored the best in most of the aspects and the difference between these options was very small and not significant. The minimum feedback scored significantly worse in most aspects than the other feedback options. And the pose acceptance of the focus image feedback was noticeably worse than the pose chain and inkspot transition feedback options.

The second subquestion asked which feedback option engaged the player the most. It seemed as if the pose chain feedback scored the best in three of the four categories and the fourth and highest category did not seem to have significant differences. However the items that should correlate within the respective categories flow and absorption, no longer did so. A possible reason could be that the adaptation of the original questionnaire violated the validation. Or the more ambiguous questions of the questionnaire were interpreted differently for an exertion game than for a sedentary game that the original questionnaire was validated with. Or the experiment was too short for the participants to actually become engaged. The inkspot transition feedback seemed to get the most attention when the participant was playing the game, because it had the highest overestimation and the largest difference between perceived and actual time. However the results were not significant. Only the minimum feedback was significantly more accurate in the time assessment than the other feedback options and thus the least engaging. Question 1B could not be given a definite answer and some more study would be required. Although a preliminary answer based on the scores would say the pose chain was the most engaging. However based on the time distortion the inkspot transition feedback seemed the most engaging.

Due to the confusion with question 1B the answer to research question 1 was based on only questions 1A and 1C. The pose chain and inkspot feedback options were the ones participants were the most satisfied with. The pose chain was also the most fun to play. Thus the pose chain feedback option provided the best user experience.

7.3.2 User performance

Subquestion 2A asked which feedback option was the most effective. Both the pose chain and inkspot transition feedback had similar results in task completion, although the pose chain had one more completed task. They were also significantly better than the focus image and minimum feedback. In pose accuracy of uncompleted tasks the inkspot transition feedback was on average more accurate than the pose chain feedback, but the results were not significant. This meant that the pose chain and inkspot transition feedback option were both the most effective. Question 2B asked which feedback option was the most efficient. There were no significant differences reported, but on average the pose chain feedback was the most efficient. The pose chain feedback was about eleven, eight and 52 seconds faster than respectively the inkspot transition, focus image and minimum feedback options in completing the task. The participants were the most quick to give up on completing the task with the inkspot transition feedback. However the sample size for uncompleted tasks with the pose chain and inkspot transition feedback were respectively two and three participants.

The small sample size was too few to accurately base results on, thus the answer to research question 2 was based on completed tasks. The pose chain and inkspot transition had a similar task completion. The difference was not significant. Efficiency showed no significant results thus only the mean was considered. The pose chain was the most efficient with about eleven seconds faster in completing a task than the inkspot transition. Thus the pose chain provided not only the best user experience, but it also had the best user performance.

Chapter 8

Recommendations for future work

This study has given a lot of answers to the research questions. However some new questions have also risen that warrant some further studies. The participants made comments about how to improve the pose game. Some of these comments were improvements on existing facets of the game. Others were added features to the game.

The pose game that was developed in this study was a fun game to play over short periods. The game was not further developed than what was necessary for the study and had room for improvement and expansion. In this chapter recommendations for continued work and new opportunities are presented.

8.1 Continuing the current work

This study focussed on the immediate feedback and its game mechanisms that guide the participant to the desired pose. The study aimed to evaluate the feedback options on their user experience and user performance. For user experience the study narrowed to pleasure, engagement and satisfaction.

The results of this study concerning engagement were a bit confusing. This could be because the participants did not become engaged or the measures were not adequate for exertion games. Further studies with another questionnaire that is more appropriate for exertion games can probably offer some better results. Nijholt et al. argued in [NvDR08] that video and audio recordings offer some more insight and physiological measurements may be more objective in determining a participant's immersion. However the sensor data should not be ambiguous, such as a motion of a limb that can either be considered input to the game or as subconscious expression. The sensors should not be intrusive either that they influence the results, such as mounting devices and wires onto the participant's body when the game only has vision based input, or devices that obstruct the range of motion when an unlimited range is desired.

It was also unclear what the effects of the exertion games were on time estimation. The experiment involved one pose per feedback for each participant which was a short trial. Maybe too short for a participant to become engaged in the game. Additional poses to assume for each feedback option may lengthen the game to allow the participants to become more engaged and more familiar with each feedback. However the single pose per feedback was an explicit choice in this study as well as not telling the details of the feedback. The experiment for the participants was kept shorter and thus getting enough participants for the experiment was more likely. Changing these choices and measures in a repeated experiment may provide stronger or different results.

This study focused on the more functional part of the feedback with a visual modality. For the pose game some aesthetic additions are definitely needed. Just having something to hear was according to some participants already a good addition to improve the experience with the game. This can be audio fragments giving hints to the desired pose, but also some background music for aesthetic value. The aesthetics in the visuals can also be improved as this study only focussed on the gameplay mechanism. For a positive influence by the aesthetics, it still requires visuals and sounds that the player likes or has a connection with [Fog03]. A study can look at what are generally accepted sounds and visuals. However it is also interesting to see if sounds and visuals can be tailored to the player. The effects on the feedback options can be studied, although aesthetics do not change the feedback's difficulty or provide more information about the pose.

The feedback in the current pose game was still simple. The results showed that the background was not a very good indicator for the distance. Participants commented that the background did not provide a clear end goal and that when they got stuck they would have liked to have some additional hints guiding them to the correct pose. A visual progress indicator, such as a bar or meter, in the foreground could take over the feedback of the color change.

Additional feedback information can be targeted at the desired pose, but most participants would prefer to know which limb or joint is wrong in their current pose. Some argued that each joint needed its own progress indicator. However one does need to question whether this would improve the user experience and user performance or that it would create more confusion. Local indicators should work with the pose processing and distance calculation of the game. The current pose game used a weighted point-to-point distance for each joint (see Appendix C). Irrelevant joints were always in the right position and thus wrong assumptions could be created.

Some additional hints could be used to improve the player's pose as well. Unlike the distance indicators these hints are only present when the player gets stuck. They could also be more ambiguous. The hints could range from anything like a sentence, image or even a number. The hints could be focused on what of the player's pose is right or wrong. It could also be about what the desired pose is. Determining when the player can be considered stuck is an interesting study to do. Which sensor data has to be considered and what kind of hint has to be given are interesting subjects to look at. Another question is what does the extra information of the feedback do with the user experience and user performance of the feedback options.

8.2 New opportunities

Mueller et al. showed with their studies that exertion games could be beneficial for social interaction [MBB10] and a social interaction could also improve one's experience with a game [Fog03]. Some of the participants thought that the pose game would benefit a lot with a multiplayer option. One participant even commented that the minimum feedback option could be really fun as a party game. A multiplayer could be realized in several ways. The first requires little change to the game. One person is trying to find a pose and other people are standing outside of the view of the camera giving the player instructions and advice on how to pose. The cooperative gameplay could even go a step further by having poses that involve multiple people. In contrast to cooperative gameplay, the multiple players could also compete against one another by assuming the correct pose before the other player does it. Further studies could look at what social interaction does to the user experience and user performance of the pose game and the feedback options.

The social factor does not necessarily have to be another player. It could be a social agent that encourages the player to continue and to persevere in order to complete the task. Such an agent could also provide further hints on the desired pose or what the player is doing wrong. The agent could range from simple onscreen instructions in text format to being an embodied agent with an animated body and a synthesized voice, such as a photographer. Studies could look at the effect of an agent on the user experience and user performance and whether the social interaction is the same with an agent as with human companions.

The pose classification process was experienced differently over the feedback options, even though it was the same with each. The pose classification could be done in several different manners. The classification could be person specific or a general classifier. The measurements could be in 2D or in 3D. A classifier could even be trained using machine learning. Studies have to be done to determine which classifier is the most appropriate for the feedback and provides the best experience and performance.

Besides a different pose classification process, some studies could be done in dynamically creating a pose descriptor and the necessary feedback resources that take the player's body proportion and the sensor's position into account (see Section 6.1). This means that instead of finding the optimal classifier the pose descriptors and the resource positions are fitted to the player when he starts to play the game.

This study only focussed on the immediate feedback during the game. But as with each game there is also a beginning and an end. The end of the game is the presentation of the reward images. The pose game that is implemented for this study has some simple reward image generation and presentation. This can be made more appealing to the player. The generation of the rewards can manipulate the image of the player and/or include props. The rewards could have a deeper meaning to them as well, which may improve the experience.

The beginning of the game involves attracting the player's attention depending on the situation. It also involves providing an explanation of what the game is about and what the player can expect. During the study this has been the task of the experimenter. This is however situation dependent. The player may be a beginner or have previous experience with the game. The game may be for leisure on a player's computer or run as a demo during an event. Attracting and instructing the player is thus given a different context.

Appendix A

Survey

In this appendix the survey for the study is presented. The items asked before and during the experiment were done in the form of a written questionnaire. After the experiment an oral interview was held.

A.1 Before the experiment

These were general questions asked to the participant before the experiment started. The purpose was to get a general description of the demographic of the participants.

1. What is your age: ______ years.
2. What is your gender: ○ Male ○ Female
3. What is your height: ______ centimeters.
4. Have you ever worked with a Microsoft Kinect Sensor: ○ Yes ○ No

A.2 During the experiment

The following questions were asked to measure the user experience for the feedback options. This questionnaire was repeated for each of the feedback options. In this questionnaire the three selected constructs of user experience were measured. The first six items measured *pleasure* and were based on the work of Bradley and Lang [BL94]. The next nine measured *engagement* and were based on the work of Brockmyer et al. [BFC+09]. The final five items gaged the *satisfaction* that the participants had with the feedback option. Depending on which feedback option the subject just did they were presented with the respective variation. At the end the subjects had the room to leave comments about the feedback option.

Indicate your overall emotional state during the session for the following emotion pairs.

1. Unhappy - Happy -4 ○—○—○—○—○—○—○—○—○ 4
2. Annoyed - Pleased -4 ○—○—○—○—○—○—○—○—○ 4
3. Unsatisfied - Satisfied -4 ○—○—○—○—○—○—○—○—○ 4
4. Melancholic - Contented -4 ○—○—○—○—○—○—○—○—○ 4
5. Despairing - Hopeful -4 ○—○—○—○—○—○—○—○—○ 4

6. **Bored - Relaxed** -4 ○—○—○—○—○—○—○—○—○ 4

Rate the items as they count for you. The score goes from *No* to *Maybe* to *Yes*.

7. I felt different from my usual everyday self. No ○—○—○—○—○ Yes
8. I got wound up. No ○—○—○—○—○ Yes
9. I felt spaced out. No ○—○—○—○—○ Yes
10. Playing seemed automatic. No ○—○—○—○—○ Yes
11. I lost track of where I was in the room. No ○—○—○—○—○ Yes
12. Playing made me feel calm. No ○—○—○—○—○ Yes
13. I really got into the game. No ○—○—○—○—○ Yes
14. I felt like I just could not stop playing. No ○—○—○—○—○ Yes
15. How long did you think you spent playing the game: ___ minutes ___ seconds.

These items are asked to measure how much each part of the feedback contributed to the experience either in the positive or the negative sense. Rate these items as how you considered them to be.

Pose chain

16. I thought my pose was accepted
    too lenient ○—○—○—○—○ too strict
17. The background was
    not useful ○—○—○—○—○ very useful
18. Seeing my image reflected was
    not useful ○—○—○—○—○ very useful
19. The notion of poses contributed
    negatively ○—○—○—○—○ positively
20. Over all this feedback was
    not enjoyable ○—○—○—○—○ very enjoyable

Inkspot transition

16. I thought my pose was accepted
    too lenient ○—○—○—○—○ too strict
17. The background was
    not useful ○—○—○—○—○ very useful
18. Seeing my image reflected was
    not useful ○—○—○—○—○ very useful
19. Searching for the pose example contributed
    negatively ○—○—○—○—○ positively
20. Over all this feedback was
    not enjoyable ○—○—○—○—○ very enjoyable

Focus image

16. I thought my pose was accepted
    too lenient ○—○—○—○—○ too strict

17. The background was
not useful ○—○—○—○—○ very useful

18. Seeing my image reflected was
not useful ○—○—○—○—○ very useful

19. The hint image to the pose contributed
negatively ○—○—○—○—○ positively

20. Over all this feedback was
not enjoyable ○—○—○—○—○ very enjoyable

Minimum

16. I thought my pose was accepted
too lenient ○—○—○—○—○ too strict

17. The background was
not useful ○—○—○—○—○ very useful

18. Seeing my image reflected was
not useful ○—○—○—○—○ very useful

19. The lack of more feedback contributed
negatively ○—○—○—○—○ positively

20. Over all this feedback was
not enjoyable ○—○—○—○—○ very enjoyable

21. Any comments about this feedback

A.3 After the experiment

After the experiment the experimenter held an interview where the following general questions were asked:

1. Rank each feedback option for how good you think they were. Which feedback option do you prefer the most

2. Why did you choose option X as the most preferable?

3. Why did you choose option Y as the least preferable?

4. What is still lacking with each feedback?

When the experimenter was unclear about answers provided by the participant, the experimenter asked him to elaborate further. The participant was also able to ask the experimenter to elaborate some more on the question.

Appendix B

Statistical results

This appendix shows the tables with the results of all the statistical analyses. The results were all calculated using SPSS 18 statistical analysis program.

B.1 User performance

	Completion frequency		ANOVA	
	Completed	Uncompleted	Mean	Std. Error
Pose Chain	26	2	0.929	0.051
Inkspot Transition	25	3	0.893	0.058
Focus Image	16	12	0.571	0.090
Minimum	10	18	0.357	0.094

Table B.1: The number of times the task has been completed or given up on for each feedback option. The estimated marginal means and standard error for each feedback option on task completion.

		Mean Difference	Std. Error	Sig.
Pose Chain	Inkspot Transition	0.036	0.058	0.546
	Focus Image	0.357*	0.107	0.003
	Minimum	0.571*	0.109	0.000
Inkspot Transition	Pose Chain	-0.036	0.058	0.546
	Focus Image	0.321*	0.109	0.007
	Minimum	0.536*	0.109	0.000
Focus Image	Pose Chain	-0.357*	0.107	0.003
	Inkspot Transition	-0.321*	0.109	0.007
	Minimum	0.214	0.111	0.066
Minimum	Pose Chain	-0.571*	0.109	0.000
	Inkspot Transition	-0.536*	0.109	0.000
	Focus Image	-0.214	0.111	0.066

Table B.2: The marginal mean differences for the feedback options based on the task completion. (*) The mean difference is significant ($p < 0,05$).

	N	Mean	Std. Deviation	Std. Error
Pose Chain	2	0.06987	0.07929	0.05607
Inkspot Transition	3	0.05502	0.05688	0.03284
Focus Image	12	0.05953	0.04690	0.01354
Minimum	18	0.23686	0.30490	0.07187

Table B.3: The mean values for each feedback option on the distance accuracy of the uncompleted tasks.

		Mean Difference	Std. Error	Sig.
Pose Chain	Inkspot Transition	0.01485	0.06498	1.000
	Focus Image	0.01034	0.05768	1.000
	Minimum	-0.16699	0.09115	0.524
Inkspot Transition	Pose Chain	-0.01485	0.06498	1.000
	Focus Image	-0.00451	0.03552	1.000
	Minimum	-0.18184	0.07901	0.185
Focus Image	Pose Chain	-0.01034	0.05768	1.000
	Inkspot Transition	0.00451	0.03552	1.000
	Minimum	-0.17733	0.07313	0.146
Minimum	Pose Chain	0.16699	0.09115	0.524
	Inkspot Transition	0.18184	0.07901	0.185
	Focus Image	0.17733	0.07313	0.146

Table B.4: The pairwise comparison for the feedback options based on the mean distance accuracy of uncompleted tasks.

	N	Mean	Std. Deviation	Std. Error
Pose Chain	26	41.206	34.544	6.775
Inkspot Transition	25	52.522	54.168	10.834
Focus Image	16	48.964	44.474	11.118
Minimum	10	92.940	82.464	26.077

Table B.5: The mean values for each feedback option on the efficiency of the completed tasks. The values are in seconds.

		Mean Difference	Std. Error	Sig.
Pose Chain	Inkspot Transition	-11.316	12.777	0.944
	Focus Image	- 7.759	13.020	0.992
	Minimum	-51.734	26.943	0.406
Inkspot Transition	Pose Chain	11.316	12.777	0.944
	Focus Image	3.557	15.524	1.000
	Minimum	-40.418	28.238	0.690
Focus Image	Pose Chain	7.759	13.020	0.992
	Inkspot Transition	- 3.557	15.524	1.000
	Minimum	-43.976	28.349	0.612
Minimum	Pose Chain	51.734	26.943	0.406
	Inkspot Transition	40.418	28.238	0.690
	Focus Image	43.976	28.349	0.612

Table B.6: The pairwise comparison for the feedback options based on the mean efficiency of completed tasks.

	N	Mean	Std. Deviation	Std. Error
Pose Chain	2	170.169	102.864	72.736
Inkspot Transition	3	136.064	113.849	65.731
Focus Image	12	159.394	58.844	16.987
Minimum	18	216.576	103.889	24.487

Table B.7: The mean values for each feedback option on the efficiency of the uncompleted tasks. The values are in seconds.

		Mean Difference	Std. Error	Sig.
Pose Chain	Inkspot Transition	34.105	83.291	0.685
	Focus Image	10.774	69.686	0.878
	Minimum	-46.408	68.007	0.500
Inkspot Transition	Pose Chain	-34.105	83.291	0.685
	Focus Image	-23.331	58.897	0.695
	Minimum	-80.512	56.899	0.167
Focus Image	Pose Chain	-10.774	69.686	0.878
	Inkspot Transition	23.331	58.897	0.695
	Minimum	-57.182	34.003	0.103
Minimum	Pose Chain	46.408	68.007	0.500
	Inkspot Transition	80.512	56.899	0.167
	Focus Image	57.182	34.003	0.103

Table B.8: The pairwise comparison for the feedback options based on the mean efficiency of uncompleted tasks.

B.2 User experience

	Ranking frequency				ANOVA	
	First	Second	Third	Last	Mean	Std. Error
Pose Chain	11	8	7	2	3.000	0.187
Inkspot Transition	8	10	7	3	2.821	0.179
Focus Image	6	8	11	3	2.607	0.171
Minimum	3	2	3	20	1.571	0.193

Table B.9: The number of times a feedback option has been given a specific rank. The estimated means and standard error for each feedback option on the rankings provided by the participants. Higher the mean is, better the rankings are.

		Mean Difference	Std. Error	Sig.
Pose Chain	Inkspot Transition	0.179	0.296	0.552
	Focus Image	0.393	0.288	0.185
	Minimum	1.429*	0.320	0.000
Inkspot Transition	Pose Chain	-0.179	0.296	0.552
	Focus Image	0.214	0.283	0.457
	Minimum	1.250*	0.305	0.000
Focus Image	Pose Chain	-0.393	0.288	0.185
	Inkspot Transition	-0.214	0.283	0.457
	Minimum	1.036*	0.296	0.002
Minimum	Pose Chain	-1.429*	0.320	0.000
	Inkspot Transition	-1.250*	0.305	0.000
	Focus Image	-1.036*	0.296	0.002

Table B.10: The marginal mean differences for the feedback options based on the rankings provided by the participants. (*) The mean difference is significant ($p < 0,05$).

	Mean	Std. Error
Pose Chain	1.250	0.248
Inkspot Transition	0.649	0.275
Focus Image	0.327	0.241
Minimum	-0.274	0.312

Table B.11: The estimated means and standard error for each feedback option on pleasure.

		Mean Difference	Std. Error	Sig.
Pose Chain	Inkspot Transition	0.601*	0.261	0.030
	Focus Image	0.923*	0.350	0.014
	Minimum	1.524*	0.297	0.000
Inkspot Transition	Pose Chain	-0.601*	0.261	0.030
	Focus Image	0.321	0.319	0.324
	Minimum	0.923*	0.246	0.001
Focus Image	Pose Chain	-0.923*	0.350	0.014
	Inkspot Transition	-0.321	0.319	0.324
	Minimum	0.601	0.322	0.074
Minimum	Pose Chain	-1.524*	0.297	0.000
	Inkspot Transition	-0.923*	0.246	0.001
	Focus Image	-0.601	0.322	0.074

Table B.12: The marginal mean differences for the feedback options based on the items scoring pleasure. (*) The mean difference is significant ($p < 0,05$).

	Mean	Std. Error
Pose Chain	3.286	0.246
Inkspot Transition	3.036	0.228
Focus Image	2.679	0.239
Minimum	2.679	0.247

Table B.13: The estimated means and standard error for each feedback option on the engagement category immersion.

		Mean Difference	Std. Error	Sig.
Pose Chain	Inkspot Transition	0.250	0.201	0.226
	Focus Image	0.607*	0.247	0.021
	Minimum	0.607*	0.234	0.016
Inkspot Transition	Pose Chain	-0.250	0.201	0.226
	Focus Image	0.357	0.253	0.172
	Minimum	0.357	0.224	0.124
Focus Image	Pose Chain	-0.607*	0.247	0.021
	Inkspot Transition	-0.357	0.253	0.172
	Minimum	0.000	0.221	1.000
Minimum	Pose Chain	-0.607*	0.234	0.016
	Inkspot Transition	-0.357	0.224	0.124
	Focus Image	0.000	0.221	1.000

Table B.14: The marginal mean differences for the feedback options based on the item scoring the category immersion for engagement. (*) The mean difference is significant ($p < 0,05$).

	Mean	Std. Error
Pose Chain	2.714	0.187
Inkspot Transition	2.250	0.190
Focus Image	2.286	0.193
Minimum	2.036	0.226

Table B.15: The estimated means and standard error for each feedback option on the engagement category presence.

		Mean Difference	Std. Error	Sig.
Pose Chain	Inkspot Transition	0.464*	0.207	0.035
	Focus Image	0.429*	0.186	0.030
	Minimum	0.679*	0.240	0.009
Inkspot Transition	Pose Chain	-0.464*	0.207	0.035
	Focus Image	-0.036	0.232	0.879
	Minimum	0.214	0.188	0.265
Focus Image	Pose Chain	-0.429*	0.186	0.030
	Inkspot Transition	0.036	0.232	0.879
	Minimum	0.250	0.278	0.378
Minimum	Pose Chain	-0.679*	0.240	0.009
	Inkspot Transition	-0.214	0.188	0.265
	Focus Image	-0.250	0.278	0.378

Table B.16: The marginal mean differences for the feedback options based on the item scoring the category presence for engagement. (*) The mean difference is significant ($p < 0,05$).

	Mean	Std. Error
Pose Chain	2.655	0.170
Inkspot Transition	2.560	0.155
Focus Image	2.298	0.164
Minimum	2.119	0.165

Table B.17: The estimated means and standard error for each feedback option on the engagement category flow.

		Mean Difference	Std. Error	Sig.
Pose Chain	Inkspot Transition	0.095	0.119	0.432
	Focus Image	0.357*	0.136	0.015
	Minimum	0.536*	0.131	0.000
Inkspot Transition	Pose Chain	-0.095	0.119	0.432
	Focus Image	0.262	0.161	0.118
	Minimum	0.440*	0.121	0.001
Focus Image	Pose Chain	-0.357*	0.136	0.015
	Inkspot Transition	-0.262	0.161	0.118
	Minimum	0.179	0.144	0.227
Minimum	Pose Chain	-0.536*	0.131	0.000
	Inkspot Transition	-0.440*	0.121	0.001
	Focus Image	-0.179	0.144	0.227

Table B.18: The marginal mean differences for the feedback options based on the items scoring the category flow for engagement. (*) The mean difference is significant ($p < 0,05$).

	Mean	Std. Error
Pose Chain	1.881	0.198
Inkspot Transition	1.952	0.204
Focus Image	1.833	0.164
Minimum	2.000	0.191

Table B.19: The estimated means and standard error for each feedback option on the engagement category absorption.

		Mean Difference	Std. Error	Sig.
Pose Chain	Inkspot Transition	-0.071	0.123	0.568
	Focus Image	0.048	0.093	0.612
	Minimum	-0.119	0.178	0.511
Inkspot Transition	Pose Chain	0.071	0.123	0.568
	Focus Image	0.119	0.115	0.311
	Minimum	-0.048	0.145	0.746
Focus Image	Pose Chain	-0.048	0.093	0.612
	Inkspot Transition	-0.119	0.115	0.311
	Minimum	-0.167	0.177	0.356
Minimum	Pose Chain	0.119	0.178	0.511
	Inkspot Transition	0.048	0.145	0.746
	Focus Image	0.167	0.177	0.356

Table B.20: The marginal mean differences for the feedback options based on the items scoring the category absorption for engagement.

	Estimation frequency		ANOVA	
	Over	Under	Mean	Std. Error
Pose Chain	21	7	1.922	0.196
Inkspot Transition	20	8	2.158	0.408
Focus Image	22	6	1.765	0.286
Minimum	15	13	1.173	0.139

Table B.21: The number of times a feedback option is over- and underestimated. And the estimated means and standard error for each feedback option on playing time estimation.

		Mean Difference	Std. Error	Sig.
Pose Chain	Inkspot Transition	-0.236	0.374	0.535
	Focus Image	0.157	0.245	0.527
	Minimum	0.749*	0.211	0.002
Inkspot Transition	Pose Chain	0.236	0.374	0.535
	Focus Image	0.393	0.487	0.428
	Minimum	0.985*	0.431	0.031
Focus Image	Pose Chain	-0.157	0.245	0.527
	Inkspot Transition	-0.393	0.487	0.428
	Minimum	0.592*	0.224	0.014
Minimum	Pose Chain	-0.749*	0.211	0.002
	Inkspot Transition	-0.985*	0.431	0.031
	Focus Image	-0.592*	0.224	0.014

Table B.22: The marginal mean differences for the feedback options based on the playing time estimation for engagement. (*) The mean difference is significant ($p < 0,05$).

	Mean	Std. Error
Pose Chain	1.047	0.181
Inkspot Transition	1.331	0.386
Focus Image	0.943	0.264
Minimum	0.433	0.112

Table B.23: The estimated means and standard error for each feedback option on perceived playing time accuracy.

		Mean Difference	Std. Error	Sig.
Pose Chain	Inkspot Transition	-0.284	0.375	0.456
	Focus Image	0.104	0.217	0.635
	Minimum	0.614*	0.194	0.004
Inkspot Transition	Pose Chain	0.284	0.375	0.456
	Focus Image	0.389	0.480	0.426
	Minimum	0.899*	0.410	0.038
Focus Image	Pose Chain	-0.104	0.217	0.635
	Inkspot Transition	-0.389	0.480	0.426
	Minimum	0.510*	0.223	0.031
Minimum	Pose Chain	-0.614*	0.194	0.004
	Inkspot Transition	-0.899*	0.410	0.038
	Focus Image	-0.510*	0.223	0.031

Table B.24: The marginal mean differences for the feedback options based on the perceived playing time accuracy for engagement. (*) The mean difference is significant ($p < 0,05$).

	Pose acceptance		Background color	
	Mean	Std. Error	Mean	Std. Error
Pose Chain	0.179**	0.150	0.429	0.279
Inkspot Transition	0.429	0.234	0.143	0.255
Focus Image	0.643	0.169	0.500	0.249
Minimum	1.071	0.181	-0.143	0.331
	Image reflected		**Pose hint**	
	Mean	Std. Error	Mean	Std. Error
Pose Chain	1.571	0.122	1.000	0.165
Inkspot Transition	1.607	0.111	1.107	0.132
Focus Image	1.357	0.179	0.821	0.191
Minimum	-0.107*	0.296	-0.964*	0.240

Table B.25: The estimated mean and standard error on mean for the 1) acceptance of the pose, 2) the color changing background, 3) the participant's image reflected on screen, and 4) the hint to the desired pose given by the feedback option. (*) Significantly different than the other feedback options. (**) Significantly different than the focus image and minimum feedback options.

	Mean	Std. Error
Pose Chain	0.679	0.173
Inkspot Transition	0.607	0.152
Focus Image	0.214	0.180
Minimum	-0.464*	0.247

Table B.26: The estimated mean and standard error on mean for the overall enjoyability of satisfaction. (*) Significantly different than the other feedback options.

Appendix C

The pose game

In this appendix the pose game will be discussed. The research is about getting people to pose in front of a camera in a specific desired pose chosen by the pose game. The pose game developed for the research is an exertion game and it uses the vision-based Kinect sensor for the player's input. The global architecture of the game is shown in figure C.1.

The feedback depends on data generated by the sensor and the data transformation process of the pose game. The camera images are processed to a pose descriptor. This descriptor is then compared to a recorded pose descriptor to measure how far the player's pose is removed from the pose that is desired by the game. If the distance has not crossed below a minimum threshold the pose game gives feedback accordingly to move the player in the desired direction.

A descriptor of a pose is a series of values for specified characteristics that describe a player's pose captured by the Kinect sensor. The game depends on a knowledge base of possible target poses, from which it selects the next target pose. For the game a number of descriptors are bundled with feedback relevant data and accumulated in the pose knowledge base.

When the player has assumed the desired pose the game will manipulate the player's image in a fictional scene as a reward. These rewards are presented to the player when all the desired poses have been assumed. The rewards are discarded when the player gives up before he has assumed all the poses.

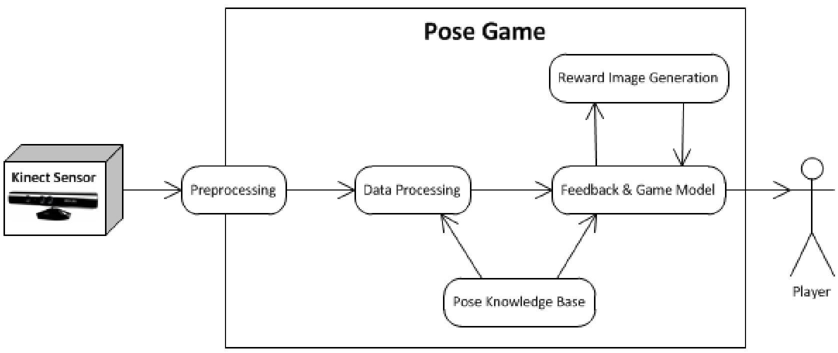


Figure C.1: The global architecture of the pose game. The Kinect sensor captures the image of the player and the pose game processes the raw data to feedback that is returned to the player.

C.1 Input devices

The input of the pose game is vision based. The Xbox360® Kinect™ sensor from Microsoft is used for the player's input. The Kinect sensor has two cameras (color and infrared) and an infrared projector. Both cameras are used to captured the scene at 30 fps on a resolution of 640x480[1]. The viewpoint of the infrared camera is translated to the viewpoint of the color camera. The communication with the Kinect sensor is implemented in the OpenNI framework[2].

For the study the pose game also has several keyboard commands implemented. These commands allow the experimenter to manipulate the flow of an experiment when necessary. These are the commands implemented in the pose game:

- Pausing and resuming a game.
- Exiting the pose game.
- Selecting a feedback order for an experiment.
- Giving up on a task in a game and continue to the next.
- Forcing a task completion in a game and continue to the next task. For debugging and determining the threshold values.

The Freeglut[3] library processes the low level keyboard events and passes them to the pose game.

C.2 Data preprocessing

The preprocessing component is handled in three stages. The first stage is done by the Kinect sensor. No special lighting conditions or a green screen background are required. The Kinect sensor can handle some basic image processing. It can register the room layout by measuring the return time of the infrared light from its infrared projector.

The second stage of the preprocessing is handled by the OpenNI framework. It retrieves the data from the Kinect sensor, which includes the raw image data from both camera's. The OpenNI framework's Scene Analyzer component provides all the labels for each pixel of the scene in the RGB camera image with either player ID's or the background (ID = 0). The OpenNI framework also generates a virtual skeleton of a tracked player with positions and orientations for each joint.

However the positions for the skeleton joints are 3D real world coordinates in millimeters with the Kinect sensor as the point of origin. In the final stage of the preprocessing the pose game transforms the position's real world coordinates to 2D projective coordinates in pixels with the top left corner of the image as the origin point and stores them in the pose descriptor. Using projective coordinates instead of the real world coordinates means that distance calculation is less effected by the size of the player compared to the size of the recorded skeleton. The player can work with the depth in order to have the same projective size.

C.3 Data processing

The data processor component of the pose game transforms the sensor data from the preprocessor into feedback data. The feedback data contains the following items:

[1] The Kinect sensor can handle 30 fps at a resolution of 640x480 and 15 fps at 1280x1024.

[2] `http://www.openni.org/`, last viewed 18-10-2011.

[3] `http://freeglut.sourceforge.net/`, last viewed 18-10-2011.

1. Distance between the player's pose and the desired pose (*real*).
2. Color value of the background (*real*).
3. Spot value for the inkspot transition feedback option (*real*).
4. Pose equality value (*boolean*).
5. Player's RGBA image (*matrix with 32-bit values*).

C.3.1 Distance between the poses

The most important task is to calculate the distance between the target pose (tp) and the player's pose (pp). This value is used to determine the color, spot and equality value and by the feedback and game model. The target pose descriptor and the relevance weights for each joint are obtained from the pose knowledge base.

The distance is calculated using equation C.1. The sum of all the separate distances for each joint is the total distance of the pose. The w parameter is the set of weights, with one for each joint. The tp parameter is the target pose descriptor and pp is the player's current pose descriptor from the preprocessor.

$$\text{TotalDistance}(w, tp, pp) = \sum_{j=0}^{n} \text{JointDistance}(w^j, tp^j, pp^j) \tag{C.1}$$

For each joint the distance is calculated with a point-to-point distance as shown in equation C.2. The position of a joint is described with 2D projective coordinates, which means that there are an x and a y value. In the equation w^j is the set of weights for joint j with separate weight values for the respective x and y values of the joint. Both tp^j and pp^j are the respective 2D coordinates with x and y values for joint j from the target and the player's pose descriptors.

$$\text{JointDistance}(w^j, tp^j, pp^j) = \sqrt{w_x^j(tp_x^j - pp_x^j)^2 + w_y^j(tp_y^j - pp_y^j)^2} \tag{C.2}$$

C.3.2 Color of the background

Based on the pose distance the game determines the feedback provided to the player. The color of the background changes from one color to another in a continuous range from 0.0 to 1.0. The color is determined by a square root curve as in equation C.3. The square root of $200 + mit$ is the maximum range of the color. In this equation the variables td and mit represent the total distance and the minimum threshold of the target pose respectively. The value of td is the result of equation C.1. The value of mit is obtained from the pose knowledge base.

$$\text{color}(td, mit) = \begin{cases} 0.0 & \text{if } td < mit \\ \frac{\sqrt{td - mit}}{\sqrt{200}} & \text{if } mit \le td \le (200 + mit) \\ 1.0 & \text{otherwise} \end{cases} \tag{C.3}$$

C.3.3 The spot value for the inkspot transition feedback

With the inkspot transition feedback the inkspot transitions to the silhouette in a stepwise progression. This means that the spot of the total distance td is within the range between the minimum threshold mit and maximum threshold mat. If the value of the spot is lower than the current inkspot's constraint the feedback will show the next inkspot in the series. Unlike the color of the background the progression of the inkspot is determined with a linear progression that is calculated with equation C.4. The value of *mat* is just like *mit* obtained from the pose knowledge base.

$$\text{spot}(td, mit, mat) = \frac{td - mit}{mat - mit} \tag{C.4}$$

C.3.4 Determining the pose equality

It still needs to be determined if the player's pose is the same as the target pose. This data is important because it decides whether the player has assumed the right pose and a reward needs to be created. Equality is calculated with equation C.5, where the result is true when the total distance *td* is smaller than the minimum threshold value *mit*.

$$\text{equality}(td, mit) = \begin{cases} true & \text{if } td < mit \\ false & \text{otherwise} \end{cases} \tag{C.5}$$

C.3.5 Image extraction

A part of the feedback itself is the player's image in the display. This is another task where vision data is transformed into feedback data. However unlike the previous data the pose descriptors are not used. The preprocessor has supplied the data processor with a RGB camera image and scene labels. The color camera produces an image in a 24-bit RGB format.

All the pixels from the camera image that have the same label as the player that is being tracked are copied to the new player's image. The pixels are copied to the same image coordinates as they are in the camera image. However the values are transformed from a 24-bit RGB value to a 32-bit RGBA value. For each pixel that is not copied a transparent color is assigned to the respective pixel in the player's image.

Although the Kinect is one of the best options for computer vision, it is still not 100% perfect. This is evident from the player's image. Around the player additional pixels from the background are also labeled as the player and thus are copied with the image data of the player. This gives the player a visible outline on the feedback (see Figures 3.1 and 3.2). The data processor does nothing to remedy this.

C.4 The pose knowledge base

The data in the knowledge base is stored in XML and picture files. In order to read XML files the pose game uses the Chilkat XML C/C++[4] library. The information in the knowledge base can be considered to exist in three levels. The highest level is the knowledge base itself, which is a collection of desired poses. The next level is the data of a single desired pose. The lowest level is a target pose descriptor with other target pose related data. A target pose descriptor is any of the poses used by the feedback options, including the first and second pose in the pose chain.

C.4.1 Desired pose data

A desired pose is all the data unique to the desired pose. It contains a list of target poses, of which the last one is the target pose for the desired pose, and additional resources, that are not explicitly connected to a target pose but relevant to the desired pose. The collection contains the following:

[4] `http://www.chilkatsoft.com/xml-library.asp`, last viewed 18-10-2011.

Target poses list: A series of specified poses that describe the movement from the base stance to the targeted desired pose. The last one in the series is used as the desired pose data. Only the *pose chain* feedback uses all of the target poses in the list.

Inkspot list: A series of pictures that are used in the *inkspot transition* feedback. These images describe a transformation of a non descriptive shape/inkspot to the silhouette of the desired pose.

Focus image data: This data contains a picture used in the *focus image* feedback. The data also contains the coordinates of where its image is drawn on screen and the size that the image needs to be drawn.

Reward pictures data: A series of picture sets (background and foreground) that are used to build the reward images for the pose game with this desired pose.

C.4.2 Target pose data

The target pose is a collection of data specific to that pose within the desired pose data. It includes all the unique information that is linked to the target pose used by the feedback options. And information that is used by the data processor in the transformation to the feedback data.

Descriptor: This is data about the characteristics of the pose, which is based on skeleton information. The data contains target values for the joints of the pose.

Descriptor weights: This is data indicating the importance of each joint of the pose. Irrelevant joints have a zero (0.0,0.0) weight and the primary joint has a weight of one (1.0,1.0). All the other joints have a weight between these extremes.

Silhouette: A picture with the silhouette information. The image is used by the pose chain feedback as an example for the player to copy.

Thresholds: A number of parts in the data processor component uses threshold values to base the feedback data on. The threshold values exist of a maximum and a minimum.

C.5 Reward image generation

The game model passes the player's image of the feedback data along with the reward picture data of the desired pose to the reward image generation component, if the game model decides that the player has assumed the desired pose. This component creates and stores the reward images for the player. The images are presented to the player at the appropriate moment.

The FreeImage[5] library is used in this component to create, manipulate and store the reward images. The current implementation of the pose game has a simple method of reward generation. The following components are used:

- The fictional scene's background.
- The player's image.
- The fictional scene's foreground.

The assumptions are made that the three images have the same size and that each image has to be copied to the origin point (0,0). First the background image is copied

[5] `http://freeimage.sourceforge.net/`, last viewed 18-10-2011.

completely to the new reward image. Next the visible pixels of the player's image are copied to their respective coordinates in the reward image. And the last stage copies the visible pixels of the foreground to their respective coordinates in the reward image.

The completed reward image is stored to the hard disk of the computer with a unique name. The component saves the path and name to the reward in a list. The FreeImage library is once more used to load the reward images using their respective paths and names and passed to the game model and feedback when the rewards have to be displayed to the player.

C.6 The game model & feedback

The game model is the central part of the entire game. The feedback is generated after the data processor has transformed the raw sensor data into feedback data that the game model can use. What feedback is displayed to the player is determined by the state that the game model is in. The low level graphics functions are handled by the Freeglut library. Not every state uses every bit of the data from the processor. Figure C.2 shows the game states and transitions in the pose game's model.

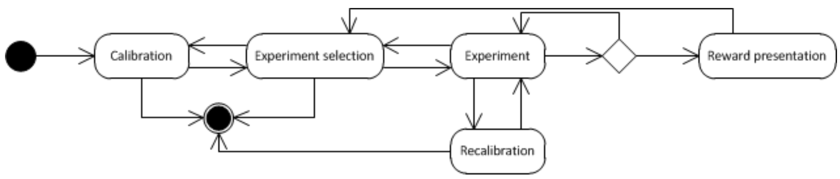


Figure C.2: This chart shows the states and transitions of the pose game.

C.6.1 Calibration

In this state the user has to calibrate a virtual skeleton for interacting with the pose game. The only feedback data that is used is the player's image, because there is no skeleton calibrated nor a desired pose selected for the data preprocessor and processor. The feedback shows every possible player in view, because there is not a single player being tracked yet.

The game can be ended in this state with a keyboard command. However once a player has calibrated a virtual skeleton the tracker gets locked on that player and the game's state transitions to the *experiment selection* state.

C.6.2 Experiment selection

This is the state of the game right before the experiment begins. The data preprocessor and processor components once more only provide the player's image as feedback data, which is displayed on screen. There is not a desired pose selected to measure the distance with. Next to the player's image is a text highlighting the four experiment groups (see Section 4.2) and which keyboard commands initialize and start the experiment.

When the player tracker gets lost this state transitions back to the *calibration* state. This state is also the second state from which the game can be ended with a keyboard command. The purpose of the game is to do the experiments and with a press of one of the appropriate keyboard commands the state transitions to the *experiment* state.

C.6.3 Experiment

In this state the full feedback data is used depending on the active feedback option. During this state the feedback data and performance data are logged by the game. With every feedback option the state updates and shows the color of the background. It also displays the player's image. And finally it checks whether the player has assumed the target pose. Some bits of the feedback is dependent on the feedback option that is active:

Minimum: This feedback option does not do anything extra with the feedback data. Neither does it display any additional visuals.

Pose chain: This feedback option displays the silhouette image that the player has to copy. It also goes to the next target pose in the chain when the player has assumed the current target pose. If there are no more additional target poses in the chain it flags that the desired pose has been reached.

Inkspot transition: This feedback option shows the abstract shape of the inkspot on the screen. It checks the spot feedback data if it has dropped below the current inkspot's constraint. And if it has crossed the constraint the feedback goes to the next inkspot in the series for the desired pose.

Focus image: This feedback shows an image of the object at its appropriate position that indicates the context. It does nothing further with the feedback data.

For this study when the tracker of the player gets lost the state will transition to the *recalibration* state. With a command from the keyboard the experiment can be ended, all the rewards are discarded and the state transitions back to the *experiment selection* state. With other keyboard commands the task can be given up or forced to be completed. A task is automatically completed when the player has assumed the desired pose. When a task is completed the game will pass the player's image and the necessary reward data of the desired pose to the reward image generation component. If there are more tasks to be completed the state transitions back to itself with a new pose to find and the next feedback option active. If there are no more tasks to be done the state transitions to the *reward presentation* state.

A fifth type of feedback option has also been implemented for this study. It is only used to determine the maximum and minimum thresholds with each target pose for in the pose knowledge base. It has the same base feedback and it shows the silhouette of the target pose. In addition, it shows the numerical pose distance in a text box. Unlike the other feedback options it does not do anything when the poses are considered equal. The player has to manually cycle through every target pose in the pose knowledge base using a keyboard command. When all the poses have been done the state considers the experiment to be stopped and returns to the *experiment selection* state.

C.6.4 Recalibration

This state is almost identical to the *calibration* state. It processes and provides the same feedback to the player. And a keyboard command can end the entire pose game. The difference is that it does not transition to the *experiment selection* state when a player calibrates his virtual skeleton and the tracker is set. Instead it returns to the same *experiment* state that transitioned to this state before. The player can then continue the experiment as if nothing has happened.

C.6.5 Reward presentation

When all the tasks are done the *experiment* state transitions to this state. The game model requests the images of all the rewards from the reward image generation component. These rewards are displayed on the screen along with the player's image. No further feedback data is provided by the data processor as there is not a desired pose selected.

The only transition from this state goes to the *experiment selection* state. This transition occurs when the respective keyboard command is given.

C.6.6 Data logging

During the *experiment* state the players activity and performance data is logged. The activity data includes the player's pose descriptor and the feedback data of each frame.

With the user performance the begin and end time of the task and the total duration of the task in milliseconds are logged. Although the time that is needed to recalibrate the virtual skeleton is subtracted from the total duration. In addition identifiers for the desired pose and active feedback are stored, as well as whether the task has been completed.

At the end of each task the logged data is saved to their respective files on the hard disk. If the experiment is stopped before all the tasks are completed the data of the last active task is discarded.

Bibliography

[And07] Mark Anders. Human joysticks: could a videogame be the secret to making our kids more fit and less fat? *ACE Fitness Matters*, 13 (5):7–9, September 2007.

[BD01] Aaron F. Bobick and James W. Davis. The recognition of human movement using temporal templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(3):257–267, 2001.

[BDDM10] Anthony J. Bonetti, Daniel G. Drury, Jerome V. Danoff, and Todd A. Miller. Comparison of acute exercise responses between conventional video gaming and isometric resistance exergaming. *Journal of Strength and Conditioning Research*, 24(7):1799–1803, 2010.

[BFC+09] Jeanne H. Brockmyer, Christine M. Fox, Kathleen A. Curtiss, Evan McBroom, Kimberly M. Burkhart, and Jacquelyn N. Pidruzny. The development of the game engagement questionnaire: A measure of engagement in video game-playing. *Journal of Experimental Social Psychology*, 45(4):624–634, 2009.

[BFH05] Petter Brandtzæg, Asbjørn Følstad, and Jan Heim. Enjoyment: Lessons from Karasek. In *Funology*, volume 3 of *Human-Computer Interaction Series*, chapter 5, pages 55–65. Springer Netherlands, 2005.

[BH05] Mark Blythe and Marc Hassenzahl. The semantics of fun: Differentiating enjoyable experiences. In *Funology*, volume 3 of *Human-Computer Interaction Series*, chapter 8, pages 91–100. Springer Netherlands, 2005.

[BL94] Margaret M. Bradley and Peter J. Lang. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1):49–59, 1994.

[BTN04] Slavador Barrera, Hiroki Takahashi, and Masayuki Nakajima. Joyfoot's cyber system: A virtual landscape walking interface device for virtual reality applications. In *2004 International Conference on Cyberworlds*, pages 286–292, 2004.

[BZ97] Richard Block and Dan Zakay. Prospective and retrospective duration judgments: A meta-analytic review. *Psychonomic Bulletin & Review*, 4:184–197, 1997.

[CK04] Anthony Chaston and Alan Kingstone. Time estimation: The effect of cortically mediated attention. *Brain and Cognition*, 55(2):286–289, 2004.

[DFAB03] Alan Dix, Janet Finlay, Gregory D. Abowd, and Russell Beale. *Human-Computer Interaction*, chapter 9 Evaluation Techniques, pages 318–364. Pearson Education, 3rd edition, 2003.

[dHGP08] Gerwin de Haan, Eric J. Griffith, and Frits H. Post. Using the Wii Balance Board as a low-cost VR interaction device. In *ACM symposium on Virtual reality software and technology*, VRST '08, pages 289–290, 2008.

[DTGP10] Liqiong Deng, Douglas E. Turner, Robert Gehling, and Brad Prince. User experience, satisfaction, and continual usage intention of it. *European Journal of Information Systems*, 19(1):60–75, 2010.

[FHVN09] Wim Fikkert, Niek Hoeijmakers, Paul Vet, and Anton Nijholt. Navigating a maze with Balance Board and Wiimote. In *Intelligent Technologies for Interactive Entertainment*, volume 9 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pages 187–192. Springer Berlin Heidelberg, 2009.

[Fik10] Wim Fikkert. *Gesture Interaction at a Distance*. PhD thesis, Universiteit Twente, Enschede, March 2010. SIKS Dissertation Series No. 2010-07.

[Fog03] Brian J. Fogg. *Persuasive technology: using computers to change what we think and do*, chapter 5: Computers as Persuasive Social Actors, pages 89–120. Morgan Kaufmann Publishers, 2003.

[GGLAC11] José-Antonio Gil-Gómez, Roberto Lloréns, Mariano Alcaniz, and Carolina Colomer. Effectiveness of a Wii Balance Board-based system (eBaViR) for balance rehabilitation: a pilot randomized clinical trial in patients with acquired brain injury. *Journal of Neuroengineering and Rehabilitation*, 8, 2011.

[Has05] Marc Hassenzahl. The thing and I: Understanding the relationship between user and product. In *Funology*, volume 3 of *Human-Computer Interaction Series*, pages 31–42. Springer Netherlands, 2005.

[HDG10] Marc Hassenzahl, Sarah Diefenbach, and Anja Göritz. Needs, affect, and interactive products - facets of user experience. *Interacting with Computers*, 22(5):353–362, 2010.

[HHTR05] Johanna Höysniemi, Perttu Hämäläinen, Laura Turkki, and Teppo Rouvi. Children's intuitive gestures in vision-based action games. *Communications of the ACM - Interaction design and children*, 48:44–50, 2005.

[HV09] Sylvain Hanneton and Anaïs Varenne. Coaching the Wii evaluation of a physical training experiment assisted by a video game. In *IEEE International Workshop on Haptic Audio Visual Environment and Games*, pages 54–57. IEEE, 2009.

[ISO98] ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDTs) – part 11: Guidance on usability, 1998.

[KMC+07] Eng Tat Khoo, Tim Merritt, Adrian Cheok, Mervyn Lian, and Kelvin Yeo. Age invaders: User studies of intergenerational computer entertainment. In *Entertainment Computing - ICEC 2007*, volume 4740 of *Lecture Notes in Computer Science*, pages 231–242, 2007.

[LCW+08] Na Li, Chun Chen, Qiang Wang, Mingli Song, Dacheng Tao, and Xuelong Li. Avatar motion control by natural body movement via camera. *Neurocomputing*, 72(1–3):648–652, 2008.

[LMW+11] Tommer Leyvand, Casey Meekhof, Yi-Chen Wei, Jian Sun, and Baining Guo. Kinect identity: Technology and experience. *Computer*, 44(4):94–96, 2011.

[MBB10] Florian Mueller and Nadia Bianchi-Berthouze. Evaluating exertion games. In *Evaluating User Experience in Games*, Human-Computer Interaction Series, pages 187–207. Springer London, 2010.

[MC96] Giovanni B. Moneta and Mihaly Csikszentmihalyi. The effect of perceived challenges and skills on the quality of subjective experience. *Journal of Personality*, 64(2):275–310, 1996.

[MDHS10] André Melzer, Ingmar Derks, Jens Heydekorn, and Georges Steffgen. Click or strike: Realistic versus standard game controls in violent video games and their effects on aggression. In *Entertainment Computing - ICEC 2010*, volume 6243 of *Lecture Notes in Computer Science*, pages 171–182. Springer Berlin Heidelberg, 2010.

[MGVA08] Florian Mueller, Martin Gibbs, Frank Vetere, and Stefan Agamanolis. Design space of networked exertion games demonstrated by a three-way physical game based on table tennis. *Computers in Entertainment (CIE) - Special Issue: Media Arts*, 6:36:1–36:31, 2008.

[MR74] Albert Mehrabian and James A. Russell. *An approach to environmental psychology.* MIT Press, 1974.

[MST+07] Florian Mueller, Gunnar Stevens, Alex Thorogood, Shannon O'Brien, and Volker Wulf. Sports over a distance. *Personal and Ubiquitous Computing*, 11(8):633–645, 2007.

[MYPVP10] Cristina Manresa-Yee, Pere Ponsa, Javier Varona, and Francisco J. Perales. User experience to improve the usability of a vision-based interface. *Interacting with Computers*, 22(6):594–605, 2010.

[NPvD+11] Anton Nijholt, Marco Pasch, Betsy van Dijk, Dennis Reidsma, and Dirk Heylen. Observations on experience and flow in movement-based interaction. In *Whole Body Interaction*, HumanComputer Interaction Series, pages 101–120. Springer-Verlag, 2011.

[NvDR08] Anton Nijholt, Betsy van Dijk, and Dennis Reidsma. Design of experience and flow in movement-based interaction. In *Motion in Games, First International Workshop*, volume 5277 of *Lecture Notes in Computer Science*, pages 166–175, 2008.

[OECN10] Miranda Occhionero, Maria José Esposito, Pier Carla Cicogna, and Giovanna Nigro. The effects of ongoing activity on time estimation in prospective remembering. *Applied Cognitive Psychology*, 24(6):774–791, 2010.

[PB10] Amanda L. Penko and Jacob E. Barkley. Motivation and physiologic responses of playing a physically interactive video game relative to a sedentary alternative in children. *Annals of Behavioral Medicine*, 39(2):162–169, 2010.

[PTF+95] Luc G. Pelletier, Kim M. Tuson, Michelle S. Fortier, Robert J. Vallerand, Nathalie M. Brière, and Marc R. Blais. Toward a new measure of intrinsic motivation, extrinsic motivation, and amotivation in sports - the sport motivation scale. *Journal of Sport & Exercise Psychology*, 17(1):35–53, 1995.

[RN08] William D. Russell and Mark Newton. Short-term psychological effects of interactive video game technology exercise on mood and attention. *Educational Technology & Society*, 11(2):294–308, 2008.

[RNRH05] Dennis Reidsma, Anton Nijholt, Rutger Rienks, and Hendri Hondorp. Interacting with a virtual rap dancer. In *Intelligent Technologies for Interactive Entertainment*, volume 3814 of *Lecture Notes in Computer Science*, pages 134–143, 2005.

[RSH+05] Liu Ren, Gregory Shakhnarovich, Jessica K. Hodgins, Hanspeter Pfister, and Paul Viola. Learning silhouette features for control of human motion. *ACM Transactions on Graphics*, 24(4):1303–1331, 2005.

[RvWP+06] Dennis Reidsma, Herwin van Welbergen, Ronald Poppe, Pieter Bos, and Anton Nijholt. Towards bi-directional dancing interaction. In *Entertainment Computing - ICEC 2006*, volume 4161 of *Lecture Notes in Computer Science*, pages 1–12, 2006.

[RZvWR06] Zsófia Ruttkay, Job Zwiers, Herwin van Welbergen, and Dennis Reidsma. Towards a reactive virtual trainer. In *Intelligent Virtual Agents*, volume 4133 of *Lecture Notes in Computer Science*, pages 292–303. Springer Berlin / Heidelberg, 2006.

[Sen05] Phoebe Sengers. The engineering of experience. In *Funology*, volume 3 of *Human-Computer Interaction Series*, pages 19–29. Springer Netherlands, 2005.

[Shi09] Dong Hee Shin. The evaluation of user experience of the virtual world in relation to extrinsic and intrinsic motivation. *International Journal of Human-Computer Interaction*, 25(6):530–553, 2009.

[SLT08] Katie Sell, Tia Lillie, and Julie Taylor. Energy expenditure during physically interactive video game playing in male college students with different playing experience. *Journal of American College Health*, 56(5):505–511, 2008.

[SMM+11] Dimitry G. Sayenko, Kei Masani, Matija Milosevic, Mark F. Robinson, Albert H. Vette, Kristiina M. Valter McConville, and Milos R. Popovic. Video game-based neuromuscular electrical stimulation system for calf muscle training: A case study. *Medical Engineering & Physics*, 33(2):249–255, 2011.

[Sol11] John Solaro. The kinect digital out-of-box experience. *Computer*, 44(6):97–99, 2011.

[TA10] Andrea Taylor and Stefan Agamanolis. Evaluation of the user experience of a standard telecare product - the personal trigger. In *Second International Conference on eHealth, Telemedicine, and Social Medicine*, pages 51–56, 2010.

[TIRC04] Peter Ulric Tse, James Intriligator, Josée Rivest, and Patrick Cavanagh. Attention and the subjective expansion of time. *Perception & Psychophysics*, 66(7):1171–1189, 2004.

[TP10] Alasdair G. Thin and Nicola Poole. Dance-based exergaming: User experience design implications for maximizing health benefits based on exercise intensity and perceived enjoyment. In *Transactions on Edutainment IV*, volume 6250 of *Lecture Notes in Computer Science*, pages 189–199, 2010.

[WMM05] Peter Wright, John McCarthy, and Lisa Meekison. Making sense of experience. In *Funology*, volume 3 of *Human-Computer Interaction Series*, pages 43–53. Springer Netherlands, 2005.

[WRK11] Jennifer R. Worley, Sharon N. Rogers, and Robert R. Kraemer. Metabolic responses to Wii Fit^{TM}video games at different game levels. *Journal of Strength and Conditioning Research*, 25(3):689–693, 2011.