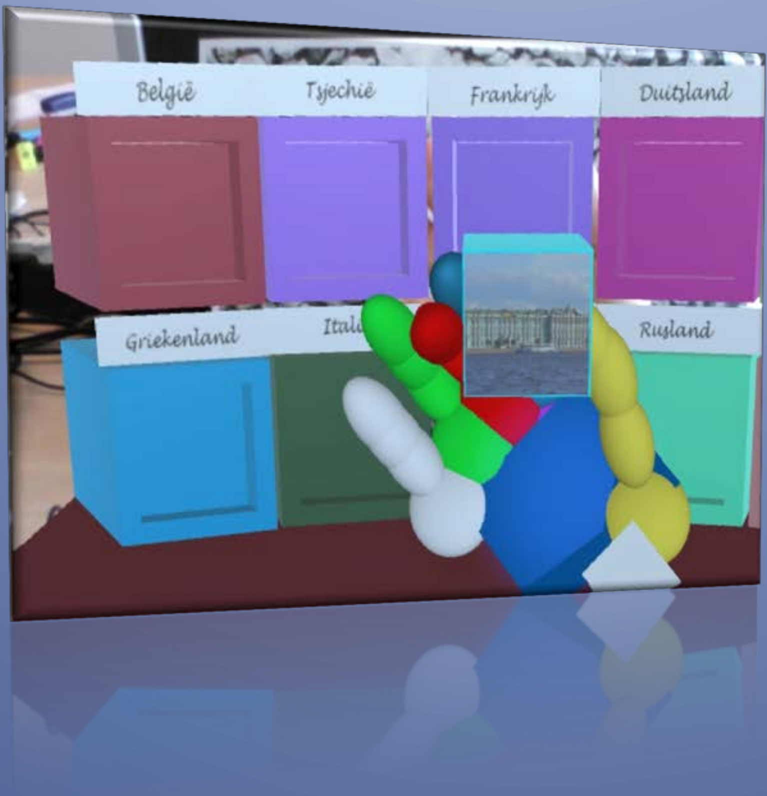
2015



Designing an engaging game for assessing motion disorders

Author: Paul Dezentje BSc. (1511165)
Daily Supervisor: Dr. Marina A. Cidota
First Supervisor: Dr. Stephan G. Lukosch
Second Supervisor: Dr. Heide K. Lukosch
External Supervisor: Dr. Paulina J.M. Bank
Committee chair: Prof.dr.ir. Alexander Verbraeck

Summary

The Technology in Motion project (TIM) has the goal to create an objective and non-invasive assessment method for motion disorders using novel technologies. Currently the TIM project has been looking at Augmented Reality (AR) as a potential technology for this purpose. A part of the project is to design a game that incites the patient to make the necessary movements, pushing him to the limit of his capabilities and so enabling the doctor to make the right decision. But a game is more than a set of steps to follow. A game that interests the player, engages him to participate, is a much stronger tool (Garris, Ahlers, and Driskell 2002; Dickey 2005) and will transform tedious repetition in a pleasant activity. But the target audience of this game will be patients with motion disorders who usually are older than 50, a group not often associated with gaming. This leads to the question "*How can an AR game that facilitates engaging motion disorder assessment be designed?*". To answer this question, a motion disorder assessment game prototype is developed and tested to find the factors that are vital to the design.

To create a game that facilitates engagement, it was decided to combine game design approaches from both serious game theories and entertainment game theories. This resulted in the game theories Rollings and Adams (2003) being integrated into the design cycle of Sebastiaan Meijer (2009). Meijer's approach was chosen because it brings a clear, iterative structure to the process and also has space to add different design theories. For these theories, Rollings and Adams provide a clear overview of important design decisions that needs to be tuned to the target audience to create the best engagement possible. The approach of Meijer is based on determining requirements to be combined with the game design theory to create a game. These requirements are the vital factors to the game design. The game is then tested and using the results of this experiment, the requirements are improved. This cycle can continue until the game is ready to be used.

To compile the initial list of requirements, two sides of the procedure were studied. First, two sessions for patients' assessment were observed and an expert on movement sciences was consulted to determine the requirements of clinicians. Second, several patients were interviewed to determine the requirements of the target audience. These two steps resulted in the requirements that led to the concept game "Post office trouble". In this game, the player has to sort packages in a sorting cabinet but instead of an address, the package only has an image on it that corresponds to a country. Using a Head Mounted Display (HMD) equipped with motion sensors will allow the game to track the hand movements of the player and use this as input. The player can thus use physical movements to play the game. The data from the sensors can be used as quantitative data for the motion disorder assessment.

The game was tested by letting eight healthy people play a prototype of the game. Two validated questionnaires measuring usability and engagement were used in the experiment. Using the data from these questionnaires it was determined if the factors found truly lead to an engaging game. The results show that the game's usability still needs improvement and that the engagement value is below average. These results can be explained in three different ways. The first option is that the engagement value could be true and the game is a little bit engaging. Secondly, the game is not engaging and the test participants were only interested because of the novelty of the AR technology. Third, the game was engaging but the low usability made it difficult for the participants to get into the game, resulting in the lower scores. All three options are possible so it is recommended that future research focusses on improving the usability so that an experiment with a larger sample-group can verify the results of this study.

Note of the author

The author used the male-centric "he" when a gender-neutral pronoun would be more appropriate. Unfortunately such a pronoun does not exist in the common English language. So, to prevent confusion only "he" was used. In the case of research participants, it was decided to also always use "he" to keep the gender of the participant anonymous.

The author wants to stress that the use of only "he" is not in any way meant as gender discrimination.

Table of Contents

Introduction

Ageing is one of the inevitable processes in the world and for many people, it comes with a problem: a motion disorder. Among people older than 65, 5.3% are afflicted with Parkinson disease, 14% have an essential tremor and these are just two examples of the many motion disorders in existence (Abdo et al. 2010). A motion disorder can be caused by neurovascular, neurodegenerative or musculoskeletal conditions resulting in a variety of symptoms which have one thing in common, they block or interfere with the normal way of motion. The amount of people with a motor dysfunction is growing (Feigin et al. 2014) and this creates a need to evaluate current assessment methods to find opportunities to help patients faster and better.

1.1. Motivation

1.1.1. Problem situation

Currently, the most common method to assess motion disorders uses qualitative visual analysis of a physical examination, often by eye. The assessment is therefore largely dependent on the experience of the doctor. There are other methods but these often use cumbersome marker-based motion capturing techniques which are avoided by the medical community (Abdo et al. 2010). The tests for movement and criteria used during an assessment session are also different for each type of motion disorder. This is necessary for better accuracy in the qualitative measurements but does create the possibility to miss other forms of motion disorders. The medical community is therefore looking for a quantitative, objective and uniform assessment method for motion disorders that is also not encumbering the patient.

1.1.2. The TIM project

The Technology in Motion project (TIM) is exploring the possibilities of developing assessment methods that use novel technologies for quantitative, cost -effective and non-invasive measurements without cumbersome apparatus (Van Hilten et al. 2014). As part of the TIM project, Delft University of Technology (TUD) and Leiden University Medical Centre (LUMC) are looking into using augmented reality (AR) to assess upper extremity motor dysfunctions. In AR the virtual content, also named the AR environment, is projected over physical reality using screens or a Head Mounted Display (HMD). The hardware and other equipment used to create the AR are defined as the AR system. In some cases the AR system includes sensors that are used to track hand motions without the use of markers, which can then be used as input to manipulate the AR environment. (Azuma et al. 2001)

The TIM project aims to use the quantitative measurements from these sensors to calculate key performance indicators of natural movements. Such an AR system can then be applied in a motion disorder assessment without encumbering the patients in their movements. Not only does this require that the AR system is capable of making accurate measurements, it also requires an AR environment that guides the patient through the necessary movements for a motion disorder assessment (Van Hilten et al. 2014). An AR system with a HMD would be ideal for this type of measurements as it can create a full 3D environment in which the patient can move (Microsoft 2015). Unfortunately, quantitative data requires multiple measurements for accuracy and the repetitiveness of the movements could become tedious for the patient, which affects the quality of the motions made. This means that the patient needs to be incited to move and somehow should not lose interest so that enough proper movements are made for the assessment. The problem that still needs to be solved is thus how to incite a patient to make the necessary movements for assessing upper extremity motor dysfunctions in an AR environment without encumbering him.

1.2. Research goal

To solve the problem, a way to motivate the patient throughout the assessment has to be developed. A solution for this can be found in the gaming industry. The hypothesis is that by disguising the assessment as interactive actions taken in a game, the patient is engaged while the assessment is performed preventing him from getting bored of the tedious exercises. Of course, a game could still be tedious and tiring but the strength of a game lays in its capability of grasping the player's attention, engaging him so that he focusses on the game and forgets everything around him (Whitton 2010). Unfortunately, changing something into a game does not automatically make it engaging (Garris, Ahlers, and Driskell 2002). To develop an engaging game requires good design in which the goal and the player's inherent interests are combined (Schell 2005). Wrong design choices can result in a game where certain elements do not match the target audience which will lead to them not wanting to play the game. The goal of this thesis is therefore to find out if it is possible to design a motion disorder assessment game for AR that is also capable of engaging the target audience.

1.3. Research questions

The goal of this thesis leads to the following research question: "How can an AR game that facilitates engaging motion disorder assessment be designed?". The idea of this question is to find the factors that are vital to the design of the game. To answer the main research question, several sub-questions need to be answered:

1. How are games designed?
2. What does engagement entail?
3. What are the opportunities and limitations of AR?
4. What is necessary to perform a motion disorder assessment?
5. How can the assessment be made engaging in a game?
6. Is the proposed game truly engaging?

The first three sub-questions focus on the necessary literature needed. First it is vital to understand how games are designed and what these processes focus on. Second, a better understanding of what engagement entails is necessary to understand what the game needs to achieve. Since the game will be in AR, its design will be subject to the limitations AR brings and can also use the opportunities that arise from this technology. These limitations need to be known as they are the boundaries of the game's design and any idea that crosses these boundaries will be useless in practice. Next, it is also important to understand more about the current motion disorder assessment methods as this is the context in which the game is developed. The most pressing matter here is to find out what actions the game needs to trigger in the player so that a proper assessment can be performed. Using all the information from the previous questions, the fifth question focusses on identifying ways to make the motion disorder assessment game engaging. Question six then focusses on verifying the results so that a proper answer to the main question can be given.

1.4. Relevance

From the literature available on game design, engagement, augmented reality and related subjects, there are several observations that can be made. First, there is a lot of research on why a game is engaging and what the effects could be of being engaged by a game(Dickey 2005; Garris, Ahlers, and Driskell 2002; Ijsselsteijn et al. 2008). The second observation is that there is a lot of research in serious game design processes often focussing on how to translate a problem situation in a game setting and how a game can then be used to analyse or solve the problem (Duke 1980; Meijer 2009). The third observation is that there is almost no research connecting these two areas. Serious game design articles talk about games being effective because they are engaging and they seemingly assume that any game designed will automatically be engaging. This thesis focuses on bridging this gap by applying engagement knowledge in altering the design process of a serious game.

Another argument for the relevance of this research is the exploration of the applicability of AR for serious gaming. With the growing interest of the commercial world into AR, research is being done in how this technology can be applied in daily life (Billinghurst, Clark, and Lee 2015). AR exists for almost 50 years in which the technology has been developed to a level in which it can be commercialized. Only in the previous decade did the research turn to exploring the possibilities for using AR. Examples are realistic training simulations, collaboration tools or design tools but there has been little about using AR in gaming. Since the goal of this thesis is to design the game using AR it creates an opportunity to explore if current technology developments are ready for AR gaming and how it can be useful in serious gaming.

1.5. Methodology

As the goal of this research is to look into the art of designing a game, it was chosen to apply Hevner's (2007) design science research cycles, which are shown in Figure 1. Hevner's theory has three parts: the environment in which the problem exist and the artefact will be used, the knowledge base comprising all the scientific knowledge and experience of the community, and the design science research which consists of the performed research. Hevner states that in order to perform good research three clearly defined cycles should overlap all three areas. The design cycle takes place in the research field and consists out of building the new artefact and evaluating it in iterative cycles. This design process also requires input and generates output. The relevance cycle collects the requirements and case information from the environment and feeds them into the research. The artefact being designed can then be field tested and will generate new requirements for the design cycle. On the other side, the rigor cycle gathers the necessary scientific knowledge to perform the research. New information learned during the design cycle is then returned to the knowledge base so that it can be used in other research.

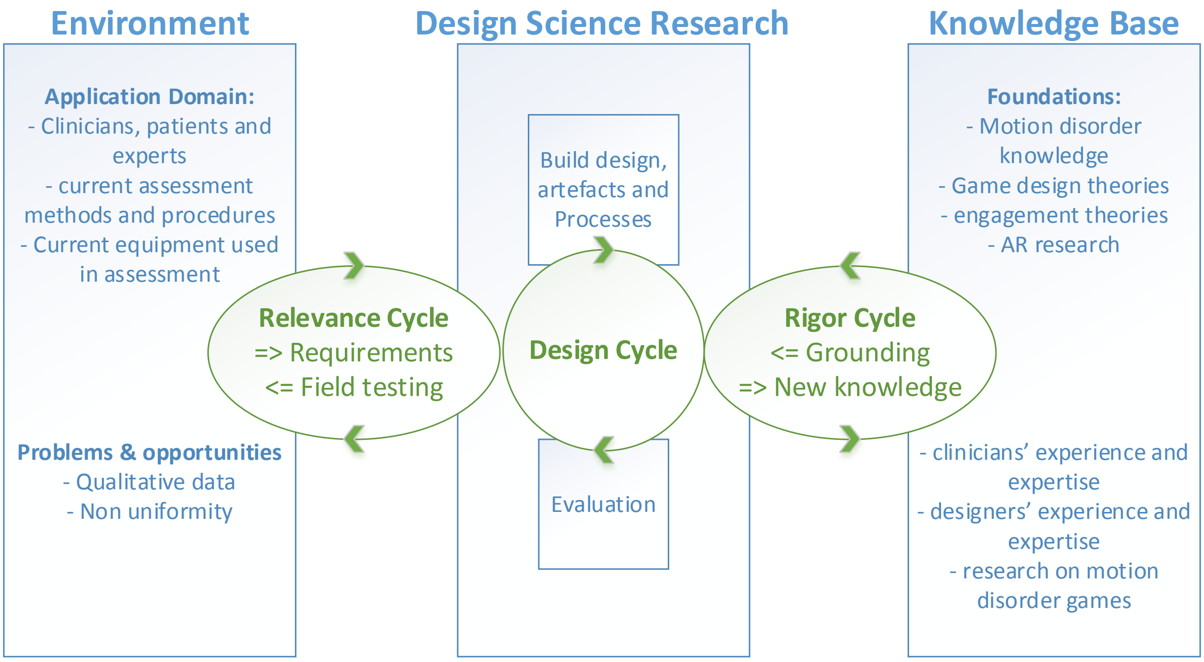


FIGURE 1 - DESIGN SCIENCE RESEARCH INTERACTION CYCLES (HEVNER, 2007)

For this thesis, the artefact designed will be an actual motion disorder assessment game. During the development, the requirements from the relevance cycle will be gathered and evaluated as these are vital factors for the game design. These requirements are thus drawn from the environment which consists of the hospital, the patients and the clinicians involved in motion disorder assessments. It also includes the current procedures used during these assessment sessions and the problem described in section 1.1. The TIM project is also part of the environment as certain directions of this thesis are dependent on choices made by this project, like the choice for AR. At the same time, theories about engagement, game design, AR and medical practices concerning motion disorders are extracted from the knowledge base through the rigor cycle. The requirements and theories are the input for the design cycle where an iterative process is used to develop the game prototype. This prototype is then field tested in the environment, again through the relevance cycle, to gather new requirements and measurements on the engagement value of the game. This information can be used to further improve the design. In the end, this will lead to a list of requirements that answers the research question. This list and other knowledge learned are described in this document and through the rigor cycle, added to the scientific community's knowledge. In the end, this will result in a prototype game with a list of important design factors that, with further development, can be used for the TIM project.

This document will first describe the related works gathered from the knowledge base in chapter 2. Then the requirement elicitation from the relevance cycle is described in chapter 3. Chapter 4 describes two iterations of the design cycle in which a concept of the game is made. This consists of developing a concept, evaluating it and then developing an improved concept. Chapter 5 describes the implementation, field testing and analysis of the developed concept, which takes place in the relevance cycle. The data gathered in this test is used to find an answer to the research question and to gather information on further improvements for the design cycle. This is then all concluded in chapter 6 which also summarises the new knowledge learned, closing the rigor cycle and finishing this thesis.

2. Related work

There are several different areas adjacent to my research question. First, there is the medical component that required knowledge about the diseases that cause human motion disorders and knowledge about current clinical assessment methods. Then, different game design techniques are researched. Engagement is also studied as this is the main feature of the game that will be measured. Lastly, a better understanding of the opportunities and limitations of AR is researched. Understanding how AR applications are different in design from applications that use other visualisation modalities is vital as one of the initial requirements from the TIM project was to use AR. Search engines like Scopus and Google scholar were used to find articles on the subjects and these were reviewed and compiled into this chapter.

2.1. Medical information on motion disorder assessment

A motion disorder is a collective indicator for multiple neurological conditions that affects the normal movements of a person. Factors that are affected are quality, fluency, speed and ease of movement. There are many forms of motion disorders; some examples are tremor, dystonia – sustained muscle contractions causing involuntary movements-, ataxia – lack of muscle movement coordination - and Parkinson Disease. (Healthcommunities.com 2015a). Another form is a motion disorder as a result of a stroke, which often results in brain damage causing motor dysfunction. The amount of people with a motion disorder is high. Among the people older than 65, 14% has an essential tremor and 5.3% has Parkinson disease (Abdo et al. 2010). The prevalence of a stroke is 11.3% in people between 60 and 79, and 29.8% in people above 80(Mozaffarian et al. 2015). Epidemiology studies on both Parkinson disease and strokes indicate that the only correlation that can be found is the age of people (Warlow 1998; de Lau and Breteler 2006) although some disorders are hereditary (Healthcommunities.com 2015a).

Some disorders can be cured with the right treatment, but in most cases the only option is to treat the symptoms, relieving pain and reducing discomfort (US National Library of Medicine 2015). There are various methods to treat motion disorder symptoms of which application would depend on the type of motion disorder. For one, there are several types of medication used for different symptoms, like antiepileptic or anti-seizure medications, Beta-blockers, dopamine agonist and tranquilizers. Although each medication is effective, they each also have possible side-effects which could bring a patient in danger (Healthcommunities.com 2015b). Sometimes medication can be ineffective. In that case surgery can be a solution. In an ablative surgery, the area of the brain that is the cause of one of the symptoms can be destroyed thus removing the symptom. Another option is a deep brain stimulation, in which a neurostimulator is implanted into the brain so that it can deliver electrical stimulation, blocking the signals that would trigger an abnormal motion (Healthcommunities.com 2015b).

As the treatment methods described focus on reducing symptoms, they require regular monitoring of the patients. These assessments are used to determine the severity of the symptoms after which the treatment can be adjusted. There are many tests that can be used for these assessments, for example using haptic feedback devices (Bardorfer et al. 2001) or motion tracking sensors (Jack et al. 2001), but as Abdo et al puts it: "*these are often expensive and time-consuming, and sometimes invasive*" (Abdo et al. 2010, 29). The common method used is a standard physical examination focused on a specific type of motion disorder in which the doctor observes and qualitatively grades, by eye and touch, the movements made by the patients on flexion, extension and fluidity of motion (Abdo et al. 2010).

The focus of the TIM project is in the assessment of the motion disorders. The goal is not so much to replace the standard physical examination but to update it with novel technology that can add quantitative data to the assessment. The game to be designed should thus support the data gathering on the same criteria. It should also be taken into account that it will be part of the treatment process of the patient and he will be playing the game every few months during a check-up.

2.2. Game design

As the goal of this project is to design a game, it is important to learn more about general game design approaches. When studying game design, there appear to be two different types of approaches, each appropriate for specific circumstances: the serious gaming approach and the entertainment game approach. In the scientific community, the focus lies mostly on serious gaming: designing a game to educate people, test policies or create public awareness of a problem. These games are usually focused on a specific issue and this is reflected in the design approaches common to serious games. These structured design approaches work by studying the issue, its environment and the goal which is then translated in a game (Hevner et al. 2004). On the other side there is the design of entertainment games. Specific design methods are harder to find here, but they are more focused on understanding the audience and how to tweak different game aspects to make the game more interesting for the players.

2.2.1. Serious game design

There are many different approaches to design a serious game. These are often very well structured, focussing on identifying the problem that the game should solve and the tools needed to solve it. Situations in which serious gaming is used include policy testing, teaching and public awareness. A good example of this is the design approach of Richard Duke (1980) which gives a clear set of nine steps to follow for the design process. The approach is shown in Figure 2. First, the specifications of the game are discussed with the game builders and the client. This includes expectations, limitations and the purpose of the game. Then the problem is expressed in a system to better understand it. The difficult nature of the problem should become clearer by defining system components and the connections between these components. The third step is to take the system and identify the components that should be included in the game and those that should be left out, thus creating a level of abstraction in the game. These components are then mapped on the twelve basic game elements identified by Duke (1980), shown in Table 1 . Each element should be expressed through one or more components. These gaming elements are then further developed and the developer adds from his repertoire to make it a game. Step six is then to construct the game and test it which is then evaluated in step seven and put out in the field in step eight.

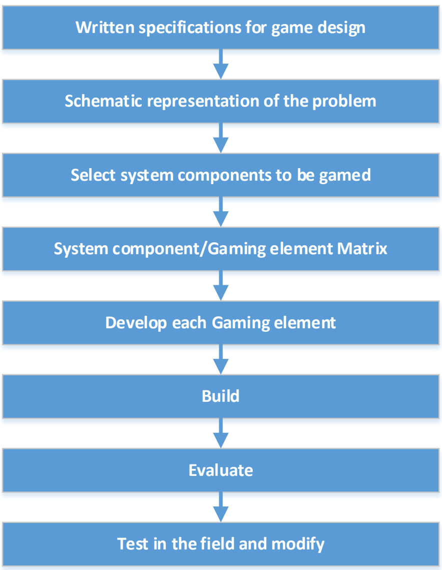


FIGURE 2 - DESIGN APPROACH OF RICHARD DUKE (1980)

Although this approach creates a good step by step procedure, the focus is mainly on solving a likely difficult problem. If this is not the case, the third and fourth step will be too specific on problem dissemination which makes the process overly complicated. Also, Duke's (1980) design process is fairly linear whereas an iterative approach would be more useful in a design process that uses new technology as new problems will appear during the design process which will require another iteration to solve.

TABLE 1 – DUKE'S (1980) GAME ELEMENTS

Nr.	Game Elements
1	Scenario
2	Pulse
3	Cycle sequence
4	Steps of play
5	Rules
6	Roles
7	Model
8	Decision sequence and Linkage
9	Accounting system
10	Indicators
11	Symbology
12	Paraphernalia

A less strict approach is the game design cycle described by Sebastiaan Meijer (2009) which is shown in Figure 3. This design cycle is a clear iterative process and gives the designer freedom to use any game design theory available. The design cycle starts with the research question (#1) and theories (#2) describing related works and tools to solve the problem. #3 constitutes describing the case environment and problems. Based on #1, #2 and #3 a list of requirements for the game (#4) is made. Next to this, information on game design theory and on published games is collected to better understand how games can be made (#5). Then a game simulation prototype is created based on #4, the requirements, and #5, the game knowledge. At the same time a test session is planned (#7) and executed #8. The data from this test session (#9) is then analysed and the list of requirements (#4) is updated. This cycle can then be iterated until the prototype is satisfactory and a finished tool can be produced (#13). Of course, a research does not have to have a finished product as result. The process itself can produce results by studying certain hypothesis. #10 and #11 therefore represent the theorized hypothesis and the induced hypothesis of the performed research.

The approaches of Duke (1980) and Meijer (2009) have a lot in common. Especially as step #6 to #9 of Meijer are the same steps as the last four of Duke. The big difference though is that Duke focusses on a systematic approach to identifying a problem, disseminating it so that it can be more easily incorporated in a game. Meijer's approach is more generalized in this matter which could make it harder to use this approach when solving a difficult problem but makes it also more accessible for people solving small problems or creating more entertainment focused games. The motion disorder assessment game's purpose is more in the direction of getting people to move while entertaining them and the design cycle of Meijer would be more applicable. As Meijer is still quite general, a game design theory fitting with the design goals of the thesis had to be found. Therefore a better look into entertainment game design was needed.

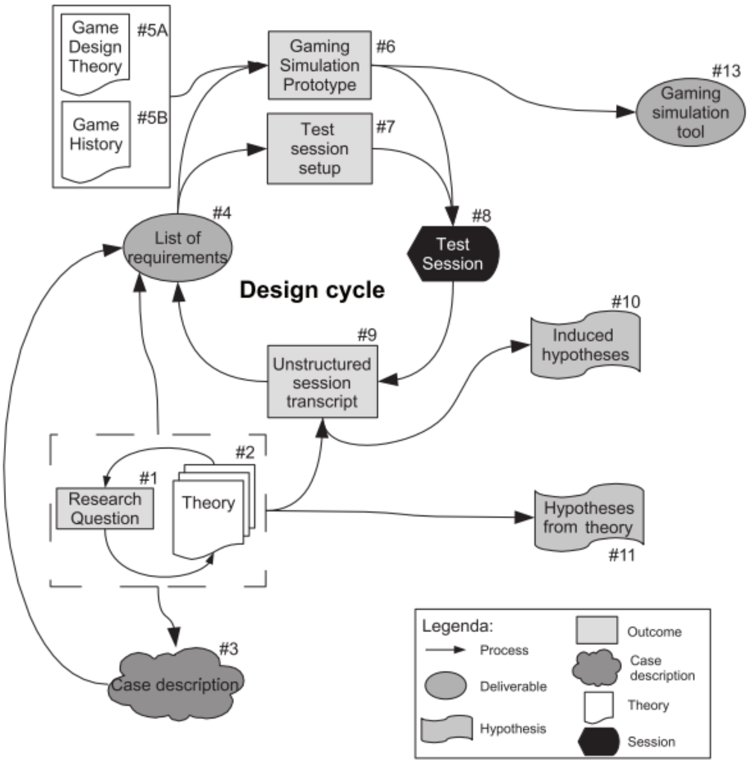


FIGURE 3 - DESIGN CYCLE BY MEIJER (2009)

2.2.2. Entertainment design

While serious game design sees a game as a tool to solve a problem, entertainment design is either market-driven or a way to express creativity (Rollings and Adams 2003). The design is therefore more focussed on potential players and how to tune the game towards them so that more games can be sold. A good approach to this type of design is found in the book of Rollings and Adams (2003) "on Game Design". The first step in their approach is to understand the audience that will play the game. By demarking for who the game is meant and researching how they react to different situations or challenges the designer creates an information basis from which design choices can be made. An important concept here is suspension of disbelieve – "A mental state in which you choose, for a period of time, to believe this pack of lies, this fiction, to be reality"(Rollings and Adams 2003, 58). As long as the player can be kept in this suspension of disbelieve, he stays immersed in the game. Breaking the suspension will disillusion the player, making him lose interest in the game.

TABLE 2 – GAME CONCEPT WORKSHEET (Rollings and Adams 2003)

#	Questions
1	What is the nature of the gameplay?
2	What is the victory condition of the game, if any?
3	What is the player's role?
4	What is the game's setting?
5	What is the player's interaction model?
6	What is the game's primary perspective?
7	What is the general structure of the game?
8	Is the game competitive, cooperative, team-based or single player?
9	Does the game have a narrative or story?
10	Does the game fall into a specific genre?
11	Why would anyone play this game?

The second step of Rollings and Adams (2003) is to take a game idea and develop it into a concept using the Game concept worksheet. This worksheet is a list of eleven questions, shown in Table 2, which covers all the important aspects of game design. The first question deals with the gameplay. A game consists out of challenges that the player faces and defining these challenge is an important step. This also includes rules that guide the actions that the player can take to face these challenges. This gameplay should then lead to some form of victory which is treated in question 2. Most games have some condition that says when the player wins. It defines what the player tries to achieve. Some games do not have a victory condition, an example being Tetris where the player goes for a high score before he loses. There is no way of winning the game as at any point somebody can still beat this high score. The third question is focused on the role the player has in the game. To make a game more interesting, the player can have a specific role like a soldier or a city planner or something more abstract. The role the player takes defines the actions he can perform in the game and should be chosen with care. When the role and the action performed do not fit together it will break the suspension of disbelieve which will result in the player losing interest in the game. The same applies to the setting discussed in the fourth question. The setting determines the content and boundaries of the game environment and should also fit to the gameplay. With the role of the player and the setting determined, the fifth question treats how the player interacts with the game world. This interaction model is different in every game and can even change during the game. Two general types of interaction models do exist: through an avatar and omnipresent. An avatar is a representation of a game character. The player can view what the avatar sees and perform the actions of this single avatar. Omnipresent means that the player can view and perform actions anywhere in the game world as if he is everywhere at once, hence omnipresent. Question six deals with the perspective of the player, how he views the game world. Examples are: through the eyes of an avatar, over the shoulder, top-down or iso-metric; which is a view from a 45 degree angle in the air. Question seven asks to describe the general structure of the game. This is done by describing the modes in the game and the relationship between the modes. A mode is a part of the game with the same gameplay. Non-interactive parts like menus and briefings are thus all separate modes. An example of a game with different interactive game modes is a game where you can both drive a car and walk around. Both the driving mode and the walking mode have their respective gameplay with separate interaction models. The relationship between the modes is that when you access a vehicle the driving mode starts and when you exit it the walking mode starts again. Question eight determines if multiple people can play the game at the same time and in what way. This significantly alters the social aspects of a game and the gameplay will be different to accommodate this change. All the information gathered above also needs to be presented to the player in the form of a story which should be explained in question nine. A story binds everything together and is the start of the suspension of disbelief. The wrong story will not interest the player and then the game ends before it was started. Question eleven asks if the game fits into a specific genre. By researching other games in this genre, tips and tricks can be found that can improve the design. The last question asks the designer to reflect on the answers given and determine if a potential player would buy and play the game.

With the game concept worksheet a better understanding of the game will be created but the game will not be ready for production. The next step of Rollings and Adams (2003) is to get feedback on the concept and improve it. Then it is time to work out all the details like the story, background information, all gameplay aspects and a complete game-structure describing step by step what actions a player can take in the game, how these actions influence the game and how this leads to the other modes in the game. This is all collected in the game script after which the building process can begin.

Rollings and Adams thus describe a process that guides the designer along the choices in the design that are important to making a successful, engaging game. The process is very specific and can thus fill in the details in the general design cycle of Meijer (2009) as the game design theory of #5 in Figure 3. Incorporating Rollings and Adams in Meijer's design cycle will require that the cycle also focusses on the target audience. Data on the target audience and their interests has to be gathered in #3. This information will then be translated into requirements. This is necessary as this information is needed to make the design choices described by Rollings and Adams, now #5 in the design cycle. These choices will lead to the creation of #6, which is the prototype.

2.3. Engagement in games

An important part of the research question is that the game should facilitate engagement. Engagement itself is a difficult term to explain. Merriam-Webster gives as definition: "To get and keep the interest of someone" (Merriam-webster 2015). This definition gives a good direction to what engagement is but it is still lacking something, it is not really clear what interest entails. Benyon et al. (2005, 61) describes engagement as follows: "being concerned with all the qualities of an experience that really pull people in" which gives a very good wording to the concept of getting someone interested. Putting them together gives the following definition: Being concerned with all the qualities of an experience that really pull people in and keeps them there.

For a game to be engaging would thus mean that the game pulls the player into the experience and keeps him there. To do this, the game needs to create the right experiences for the player. Engaging someone in a game is difficult as you have to tune it to a person's preferences (Whitton 2010). Of course, each individual is different thus it is important to find general preferences of the target audience. IJsselsteijn et al. (2007) also mentions the importance of understanding the motivations of the elderly, which overlaps with the target audience of the motion disorder game, but notes that there is a gap in scientific knowledge about this, arguing that further research on these motivations is necessary. The question that remains is how to further delineate engagement.

Ijsselsteijn et al.(2008), while developing a questionnaire to measure engagement, studied game experiences and separated them in seven different components that contribute to engaging the player. These components, shown in Table 3, are focused on the feelings a person experience during play like how competent he feels or how the game allows him to be imaginative. Flow here stands for "the state in which people are so involved in an activity that nothing else seems to matter; the experience itself is so enjoyable that people will do it even at great cost, for the sheer sake of doing it"(Csikszentmihalyi 1992, 4). The other feelings are how challenged the player feels, general positive feeling, general negative feeling like being bored and feeling that the game itself is irritating or annoying (Ijsselsteijn et al. 2008).

TABLE 3 - IJSSELSTEIJN ET AL. (2008) GAME EXPERIENCES

#	Experiences
1	Competence
2	Sensory and imaginative immersion
3	Flow
4	Challenge
5	Positive affect
6	Negative affect
7	Tension and annoyance

To create an engaging experience, the player should feel the first five feelings and not feel the last two. Although this gives a better understanding of what engagement looks like, it helps little in the way of game design. A different way is described by Jesse Schell (2005) who says that the interest someone shows can be explained with three factors. First there is the inherent interest each person has in specific topics, second there is the way you present it and lastly is the psychological proximity.

Psychological proximity means that if something happens close to you, it automatically is more interesting or important. In a game, this can be created through empathy. By creating an accessible character the player can empathize with, the psychological proximity can be made smaller and this makes the character more interesting. If the interest gathered from these three factors can be kept, the player is engaged.

Michele Dickey (2005) takes a different angle on explaining engagement. She says that to create engagement a strategy has to be developed that fits game elements to the players preferences. The game elements she uses are Rollings' and Adams' (2003) five game dimensions explained in Table 4. Deciding how to use each dimension is an important step in developing an engaging game as each choice should be carefully made to fit the target audience. Engagement here has some similarities to Rollings' and Adams' (2003) suspension of disbelieve. It can even be said that suspension of disbelieve is caused by the player being engaged by the game, thus not noticing reality. If a choice in one of the dimensions is so farfetched that the player cannot believe it anymore, he realizes it is a game and is not engaged by it anymore. Thus the suspension of disbelieve is also gone.

Engagement is a very difficult topic and all theories described are more complementing each other than being mutually exclusive. Rollings' and Adams' (2003) game dimensions are the area in which the design choices are made. These choices have an effect on the different feelings of the player as described by Ijsselsteijn et al. (2008). Depending on the inherent interest of the player, presentation, and the psychological proximity as described by Schell (2005), a design choice has a certain effect on the feelings that affect engagement.

TABLE 4 - THE FIVE GAME DIMENSIONS (Rollings and Adams 2003)

Dimension	Explanation
The Physical dimension	The space in which the game takes place. This consists of dimensionality, scale and the boundaries of the space.
The Temporal dimension	The role of time in the game. Is there a continuous or discrete time flow? Is time anomalous? Can the player adjust time?
The Environmental dimension	The appearance, atmosphere and cultural context in the game.
The Emotional dimension	The emotions of the characters and the emotions intended to be invoked in the player
The Ethical dimension	The moral and ethics that consist in the game-world and the way these are presented to the player.

2.4. Augmented Reality

One of the predefined requirements of the project is to design the game for an AR environment which is viewed through an HMD. A common definition of AR is given by Ronald Azuma in three requirements (Azuma 1997):

1. AR combines real and virtual content
2. AR is interactive in real time
3. AR is registered in 3D

AR thus adds interactive, 3D virtual elements to the real world to improve its value. This could be by adding entertainment, information, or communication tools. Examples of new technologies that augment reality are Google glass (Gibbs 2014) and Microsoft Hololens (Microsoft 2015), displaying relevant information over reality, for example as in Figure 4 . Both of these devices are HMDs but AR can also be used on a flat screen. Applications like Layar (Layar 2015) add information over a live camera feed, thus creating an AR on a mobile device.

FIGURE 4 - CONCEPT IMAGE OF AR VIEW (Microsoft 2015)

AR already exists for more than 50 years but is becoming more popular as recent technological advances allow application in daily life. The development of AR was especially supported by the military. Starting in the 1960's, the US air force researched various ways to support their pilots with AR and even now, companies provide AR solutions for military purposes (Billinghurst, Clark, and Lee 2015). In the 90's, companies started researching other uses of AR and continued research in tracking, display techniques and interaction. At the same time, advances in wearable computing and microcomputing created new opportunities to use AR and, starting around 2007, the first commercial purposed AR systems were available (Billinghurst, Clark, and Lee 2015).

AR is thus at the starting point of being used in various fields. Most research is therefore still exploring the possibilities that this upcoming technology brings and studying these works can help identify the possibilities that will be of use for the motion disorder assessment game. The next paragraphs will therefore explain a variety of fields in which AR is applied and what information can be useful from these studies.

First of all, AR can be used as a collaboration tool. Shen et al. (2010) describes a system in which a multi-disciplinary team can observe, explore and modify a product design in a 3D environment using HMD's. They also note that the use of visual clues (e.g. highlighting features) greatly improves communication between users. A different collaboration example is described by Poelman et al. (2012) in which a HMD system is used by crime scene investigators in the field to collaborate with experts at base. To facilitate the collaboration, the system also offers a toolset with virtual tools for marking, measuring and capturing data. A problem they did find in collaboration is that an interaction protocol is needed to prevent communication chaos. It is also important that the investigator is able to see the experts to give him the feeling that he is part of a team and is not being observed from above. The collaboration techniques used can also be useful for the clinicians. It will allow easy communication between doctors at different hospitals and add the possibility for a second clinician to observe and improve the assessment.

A different way to use augmented reality is to help people with phobias by gradually exposing a patient to his fear in a safe environment (Wiederhold et al. 2002; Botella et al. 2005). Botella et al. (2005) developed a system in which people with a cockroach phobia are shown cockroaches moving through the room. The power of this treatment is not just the image. By letting the cockroaches move realistically through the room and allowing them to be killed if the patient touches them, the virtual elements allow the player to be engaged by the environment, enhancing the effect of the treatment. This research shows the importance of engagement as the study shows that viewing an image of the cockroach would have lesser effect (Botella et al. 2005).

Another example of AR usage in the medical world is explained by Khademi et al. (2012). They describe a system that combines AR and a haptic hand-held device to measure human arm's stiffness. The system can help improve the rehabilitation of patients because it would allow them to train without a therapist. The system uses visual sensors and the haptic device to measure arm movements, which are then displayed on a screen. The haptic device is used to create more accurate measurements which at the time were not yet possible with the visual sensors. Of note in this research is that they found that visual tracking software lacks the accuracy needed for their tests. Since these tests are very close to what the TIM project tries to achieve, it should be taken into account that tracking is still in development (Billinghurst, Clark, and Lee 2015) and that the possibilities are still limited to what can be achieved with it.

Alamri et al. (2010) also works in the same direction and developed a framework to be used when developing an AR – haptics rehabilitation game. In this framework they also describe two exercises: a shelve exercise where the patient has to remove a virtual object from a shelve and place it somewhere else, and a cup exercise where the patients has to follow a specific path with the cup. Both exercises measure normal daily activities but the reason for this choice is not described in the article. Alamri et al. (2010) also says that it is important to measure the time it takes to complete every task. They measure the task-completion time of horizontal movement, vertical movement and following the guiding axis. This framework and especially the exercises are close to the purpose of the motion disorder assessment game and give a possible direction to the game design.

All the articles described till now talk about rehabilitation exercises whereas the TIM project wants to use these novel technologies for assessments. Here, only Broeren et al. (2002) uses a haptic feedback device to assess upper extremity motor dysfunction as a result of a stroke. They show that their AR system has potential to be used for clinical assessments, but the haptic feedback device used is a good example of the cumbersome motion capturing techniques that the TIM project tries to avoid.

2.5. Summary

Motion disorders are neurological conditions that affect the normal movements of a patient. Treatment often consists of reducing the symptoms as curing it is impossible. This requires regular assessment to keep track of a patient's progress. These assessments consist of a physical examination focused on the type of motion disorder as other options are not worth the expense. The TIM project thus works on creating new assessment methods that are less expensive and risky than the other methods and also work better than a standard physical examination.

The goal of this research is to design a game to support this assessment method but there seem to be two views on game design. From the serious game design view, the game is designed as a tool to solve a problem and it focusses on identifying and incorporating the problem and possible solutions in the game. On the other hand entertainment game design is either market-driven or art-driven. This often translates in a focus on potential players and how to make the game interesting for them. The motion disorder game could be categorised as something in between these two types of games. The goal and function of the game is a serious game with the intention of gathering data for medical analysis. On the other hand, the patients play the game as if it is an entertainment game. The game should entertain them while the measurements take place in the background.

Therefore a design approach is used that incorporates Rollings and Adams (2003) game design theory into the game design cycle of Meijer (2009). This does not change the procedure in Meijer's (2009) design cycle but instead some of the points will focus on gathering information needed for the design choices described by Rollings and Adams (2003). First of all, the case description, #3 in Figure 3, will include understanding the inherent interest and preferences of the players. These are then translated into player-specific requirements for better engagement which are added to # 4. The design cycle left space to choose the game design theory in #5a for which the theory and practices of Rollings and Adams will be used as these give a good framework of the choices that have to be made to create an engaging game. The choices are then made based on the complete package of requirements in #4 to create the prototype in #6. After testing this prototype in #7 and #8, #9 is to analyse the data to find new requirements for the game. This step again has an extra focus on the player-specific needs so that, in the next iteration, the design choices can be better tuned.

During the design process, it is important to design the game so that it is engaging for the player. Engagement means being concerned with all the qualities of an experience that really pull people in and keeps them there. An engaging game will thus keep the player entertained and should make measurements easier to take. To make a game engaging the players' inherent interest and preferences in presentation and psychological proximity have to be known. Game design choices have to be made in such a way that they will positively influence the experiences of the player, thus creating the engagement.

The game design cycle in Figure 3 also has a point about game history, 5b. The TIM project is developing the assessment system using augmented reality (AR). Studies show that using AR can be helpful for the rehabilitation of motion disorder patients. It can also be used for the assessment of motion disorders but there have not been many studies in this area. The key-point in using games in AR is to create engagement by using inherent interests and the expected behaviour of the virtual environment to draw the player into the game. Alamri et al.'s (2010) framework shows simple exercises that give a direction to a possible motion disorder assessment game. An additional function that could improve the game is a collaboration tool that allows easier exchange of information between clinicians for improved assessments.

3. Requirements elicitation

Before the game can be designed, a clear list of requirements needs to be created. These requirements are based on the current system that is used by clinicians, the gaps in this system and the wishes from different actors.

3.1. Procedures for compiling a list of requirements

To examine the current situation, the case is studied from two sides: the experts who perform the assessments and the patients who are being assessed. To find the requirements of the clinicians, two motion disorder assessments were observed at the LUMC and experts where consulted to understand the current procedures and what could be improved. As the goal of this research is to design an engaging game, it is also important to better understand the players of the game, namely the patients. Therefore interviews are conducted to determine what they would require from the game.

3.1.1. Actors involved

For the motion disorder game, two groups of actors can be distinguished. First there are the clinicians who will use the system to perform the assessments. Because a motor dysfunction can be the result of various problems these clinicians come from different fields like neurology or rehabilitation (Abdo et al. 2010). Next to the attending clinicians there are also motion science researchers who have a lot of knowledge about various motor dysfunctions and will be using the system as well.

The second group of actors important for the design process are the future players of the game, namely the patients. To be able to engage the target audience it is important to understand what their needs, interests and issues are. In general, motion disorder patients are older than 50 years of age (Feigin et al. 2014). Other than age, there are no other social or natural distinctions in the target audience. This means that when doing tests the sample group should be chosen carefully so that, except for age, personal characteristics are spread out so not to bias the results.

3.1.2. Observations and consulting experts

The primary goal of the game is its function as an assessment tool for different motion disorders. As the system needs to improve the quality of assessments, it is first important to understand how current assessments take place and why there is need for improvements. Therefore, two assessments where observed at the LUMC. The assessments chosen were for two different disorders, Parkinson Disease and a stroke, to create a broader view of the issues that the system developed by the TIM project should address. During these assessment sessions, the researcher was in the room with the patient and the clinician. The clinician performed the assessment as usual while explaining his thoughts and reasons behind each step made. After the assessment the researcher was given the opportunity to ask questions to both the patient and, after the patient left, the clinician. During these observations only notes were made out of respect of the patient's privacy. The notes of both observations are then analysed by eliciting each requirement for the assessment. These are then compared to determine which major requirements are needed for an assessment.

To further explain the current procedures and issues, the second study involved consulting an expert in movement sciences from the LUMC. This consult was planned after the observations so that these could be discussed as well. The consult will focus on discussing questions found in the observations and to understand which movements are required to be incorporated in the game. The answers given by the expert are then compiled in a set of requirements for the game.

3.1.3. Interviews with patients

To understand how to engage the target audience it was decided to perform interviews to determine what they would require of a game. It was chosen to use interviews because this would allow the patients to better express their opinions as opposed to a survey. It will make it easier to understand what their preferences are and, as the goal was to get subjective data, there is also no problem of interviewees unintentionally changing facts. Also, a small number of good interviews are preferable to a large number of small bits of information gathered from a large-scale survey. It was chosen to focus the interviews on Parkinson patients as this would allow the researcher to perform the study at one department of the LUMC, namely neurology. Also, Parkinson disease can happen to anyone and happens mostly to people older than 50 (Stichting Parkinsonfonds 2015) so the sample group is similar to the target audience. The focus of these interviews was mostly to create a starting point for the design process which could later be refined through multiple iterations of the design cycle.

Seven semi-structured interviews were held with Parkinson patients at the neurology department of the LUMC. Patients were asked by the attending clinician if they would like to participate in the study. Participants were received in a separate room in which the interview of about 30 to 45 minutes took place. First, a small presentation of the project was given. Then the interview was performed and was recorded to allow for better analysis of the answers.

TABLE 5 – SEMI-STRUCTURED INTERVIEW FOR REQUIREMENT ELICITATION

#	Question
1	What is your age?
2	What is or was your occupation?
3	What is your computer experience?
4	What is your computer game experience?
5	What is your analog game experience?
6	Do you prefer a realistic or a fictional setting?
7	How should cultural rules and morals be treated?
8	Which art style do you prefer?
9	Would you prefer time sensitive exercises in the game?

The interview, shown in Table 5, was divided in two parts. The first five questions focused on the gaming history of the patient and the reasons for the choices he made in this regard. The second part is based on Rollings' and Adams' (2003) five game dimensions but do not go in depth. Instead, they ask for the information needed to make the design choices within these dimensions. Question 8 refers to art-styles which was be answered by the interviewees by picking one of the four pictures shown in Figure 5 and explaining why this option is preferred.

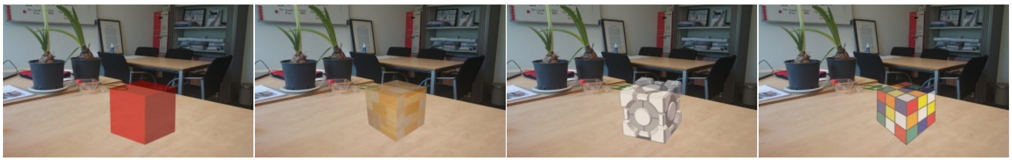


FIGURE 5 - ART STYLE CONCEPTS

The interviews resulted in audio recordings of the conversations. These recordings where then transcribed into bullet points describing the core of the message. Per question, the answers were compared to identify common trends. If more than half of the participants supported a trend it was noted and translated in a requirement for the game. If the opinions were divided among multiple trends, the one most often mentioned was chosen as the requirement but the others were also noted in case a later iteration of the design process discredits this requirement.

3.2. Results of observations and discussions with experts

Both of the observed assessments had a similar setup. First the doctor asked the patient how he was doing, giving him time to tell his own observations. Then the doctor asked the patient to make several movements and observed these movements using a checklist with Likert scales as an analysis tool. Using this information the clinician then made an analysis of the situation and adjusted the treatment of the patient in question. In the case of the stroke patient, the movements made were focused on identifying the range of movements the limbs were able to perform. The patient was asked to make several movements that tested each joint to its limits. In the cases that this was very difficult, the doctor actively moved the limbs to feel the tension and resistance in the joints.

The assessment of the Parkinson patient was focused on identifying the speed, accuracy and amplitude of movements. Exercises like finger tapping on a table, repeatedly touching the thumb and index finger together and heel tapping were used to identify the severity of the Parkinson disease. Then the patient was asked to perform several movements to identify if the limbs froze before a motion was made. The patient was also pulled out of balance to test how quick he can stabilize again. Lastly the patient had to hold his hands in three positions, a static position, a rest position and in motion, while the doctor determines if there is a tremor visible.

In the discussion afterwards, both doctors indicated that a usual session takes between 30 and 60 minutes depending on the severity of the case. The doctors also agreed that each type of assessment uses different motions which are specific for the symptoms of that motor dysfunction. The consult with a movement science expert[1] therefore focused on identifying which movements could be used in an assessment method meant for all the different types of motor dysfunction. The expert said that it would be best to focus on functional movements as these are motions that are used in everyday life. These movements are not aimed at specific symptoms. Instead it can be analysed what type of hinder the patient experiences during his normal routine. The expert said that to perform a proper assessment the five motions shown in Table 6 should be measured.

TABLE 6 - FUNCTIONAL MOVEMENTS

#	Movement
1	Range of motion in individual joints
2	Reachable workspace
3	Pointing and reaching
4	Reach-and-grasp
5	Tremor assessment

[1] Dr. P.J.M. Bank at the Leiden University Medical Centre

The range of motion in individual joints means that the maximum flexion and extension in the fingers, wrist, elbow and shoulder should be examined. The reachable workspace is the space that the patient can reach by only moving the hands and arms and is determined for each arm. An image showing an example workspace of a person is shown in Figure 6. Measuring pointing and reaching is done overhand and underhand as well as on the ipsilateral side and the contralateral side. The same applies for reaching and grasping plus that the patient has to perform several grips at various distances from the body. The tremor assessment is about measuring the frequency and amplitude of a possible tremor in the hands. This should be done during the mentioned functional movements and in a resting position. When the five types of measurements in Table 6 are combined in the game it will create enough input for an assessment.

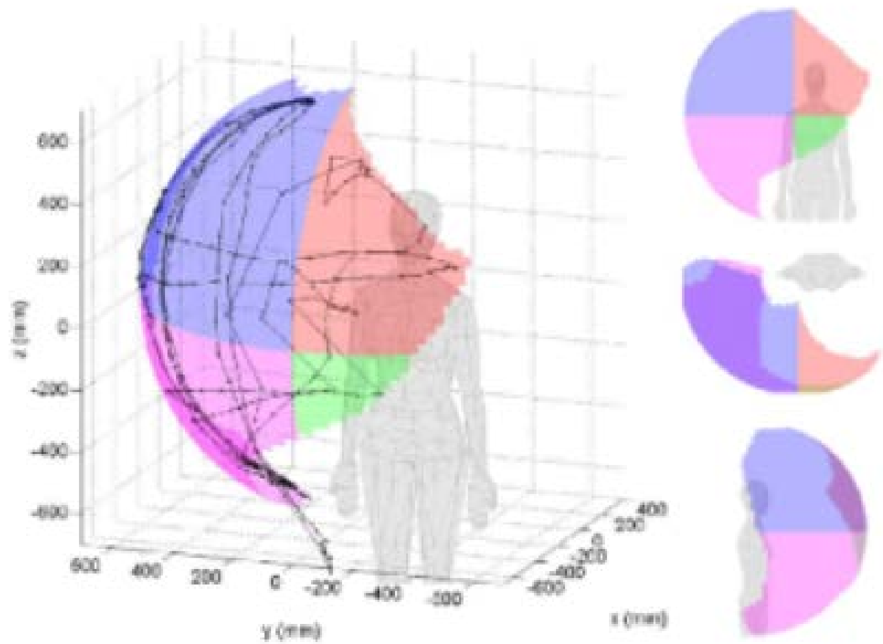


FIGURE 6 - EXAMPLE OF REACHABLE WORKSPACE (Han et al. 2013)

3.3. Results of interviews with patients

The interviews were performed with seven Parkinson patients; the raw data is available in Appendix A. The personal information they have given show ages between 57 and 80 and a variety of jobs from political journalist to car mechanic. When asked what their experience with computers was all of them indicated that they saw the rise of the computer age in the latter part of their careers and started using them for administrative tasks. All of them still use the computer but mostly to check e-mail or the internet. When asked about the experience with games, both digital and analogue, six indicated to still play analogue games and four also played digital games. The reason they indicated for playing games is twofold. First of all they like the social elements involved in the games, enhancing their interaction with friends and family. Secondly, all interviewees mentioned that they like to play games to keep practicing their mind and memory, keeping their brain trained. They also show a clear dislike for any form of violence in games. When discussing the game dimensions six of them indicated that they prefer a game that is close to reality. Five of the participants also indicated that they do not like time pressure in a game, preferring to take their time to think things through. They also indicated a preference for a more cartoonish form of art, the third option in Figure 5, to be used in the game. When discussing the usage of different cultural rules they also indicated that things should stay realistic but also that it would not be a problem if things would be simplified somewhat for the sake of the game.

3.4. Determining important requirements for the game

Based on the results of the observations, interviews and consults, certain requirements that the motor dysfunction assessment game should incorporate can be identified. Firstly, the observed assessments showed us that the clinician wants the patient to make movements to the limit of his capabilities. Only by measuring the limits of the patient's movements can a proper assessment be made and this is thus the first requirement of the game:

R1: The game needs to incite players to move at the limit of his motion capabilities.

The two observations also give a good example of the gap that the TIM project wants to fill. Both doctors use the same type of methods, although with different movements, and analyse these movements with qualitative methods. Prime example of this was the checklist used during the assessments, were a Likert-scale was used for each movement and that all these measurements are based on the sight or touch of the doctor. If the system to be developed is going to be used for different types of motion disorders more accurate measurements will be needed and these measurements should be quantitative instead of qualitative.

R2: The game needs to be supportive of quantitative data retrieval.

Another observation was that the time both assessments took was between 30 to 60 minutes. As the game is a supporting tool, it should not add more time but should instead only replace the time it took for the doctor to make observations. This means that the game will be played for a maximum time of 30 minutes so that the clinician has enough time to talk with and assess the condition of the patient. Another reason for this time limit is that the patient will get tired the longer he plays, decreasing physical abilities and affecting the measurements. This could also be considered as part of the patient's capabilities and the game should be able to measure this. Therefore it would be important that it is possible to adjust the duration of the game. This would also make it easier for the clinician to stop the measurements when enough data is collected or continue measuring if necessary.

R3: The game must not be longer than 30 minutes.
R4: The game's duration needs to be easily adjustable.

During the consult with the expert in movement sciences it was decided to focus on functional movements. Implementing all will be hard as a game will focus on one or two specific types of input. Still, as many of the five functional movements shown in Table 6 should be incorporated in the game.

R5: The game needs to incorporate as many functional movements as possible

The interviews with the patients show that the target group is not very accustomed to computers and especially digital games. Therefore it is prudent that the system is easy to use and has an intuitive interface.

R6: The game has to have an easy to use, intuitive interface

Furthermore the interviewees played games for social interaction and to train their minds. A social element can be very difficult because an assessment is a private affair. The game should thus focus on creating brain exercises that challenge the player. At the same time, there is a large variety in the education, interests and social background of the patients which means that the game should be accessible for everyone older than 50.

R7: The game's exercises need to challenge the mind of the player.
R8: The game needs to be accessible for a large variety of backgrounds.

In the interviews, the patients also gave their preferences in certain design choices. They prefer a realistic setting in the game with a cartoon art style. They also prefer games without any form of time pressure.

R9: The game has to have a realistic setting with a cartoonish art style.
R10: The game should have no form of time pressure.

3.5. List of requirements

Table 7 shows the requirements that where identified in the experiments described in chapter 3. By listing the requirements in the table, it also shows that there are some difficult design choices to be made where two or more requirements are in conflict with each other. Requirement R7 and R8 show such a design choice as a challenging game for one person can be easy for somebody else. This could be because of a difference in education level but also because of a difference of interests. The product should be designed in such a way that both requirements are met. Another design issue is that the game should take less than 30 minutes (R3) and should have no time limits (R10). It could be that this time limit is too short to make good measurements (R2) without pushing the patient with time pressure. When designing the game this issue should also not be forgotten.

TABLE 7 - FIRST SET OF REQUIREMENTS

Nr.	Requirement
R1	The game needs to incite players to move at the limit of their motion capabilities
R2	The game needs to be supportive of quantitative data retrieval
R3	The game must not be longer than 30 minutes.
R4	The game's duration needs to be easily adjustable
R5	The game needs to incorporate as many functional movements as possible
R6	The game has to have an easy to use, intuitive interface
R7	The game's exercises need to challenge the mind of the player
R8	The game needs to be accessible for a large variety of backgrounds
R9	The game has to have a realistic setting with a cartoonish art style
R10	The game should have no form of time pressure

4. Designing a game for assessment of upper extremity motion dysfunction

With the information gathered in the previous chapters the concept of the upper extremity motor dysfunction assessment game can be created. This concept is designed in two iterations of Meijer's (2009) design cycle. First, an initial concept is created which is reviewed by experts on game design. These reviews start the second iteration in which the concept is improved and worked out in further detail.

4.1. Game concept

Following the design cycle from Figure 3, a game is designed by combining #4, the list of requirements gathered in chapter 3, with #5, the game design theory and history studied in chapter 2. Combining this information gives all the tools needed for a first concept design of the game.

4.1.1. Procedures and methods used

First, multiple possible concept ideas based on the list of requirements from Table 7 were developed. If an idea was in obvious conflict with one of the requirements it was dropped until only promising ideas where left. These ideas are then compared to each other and the one who best fits to the requirements is developed further. Following Meijer's design cycle, the design theories (#5) from Rollings and Adams (2003, 53), which are described in chapter 2.2, are applied on the chosen concept. Thus the game concept worksheet is used to further delineate the concept using it as a checklist to guide the process along all the important aspects of the game design.

4.1.2. First concept ideas

Three concept ideas that looked promising for the motor dysfunction assessment game were developed. The first idea is to create a railroad setting in which the player has to clean up the rails from an omnipresent perspective. Using AR, a table is transformed in a country landscape with trains connecting the villages. Obstructions on the rails are constantly created and the player has to remove them to allow the train to pass. The obstructions will come in various shapes and sizes to train different grasp and reach movements.

The second idea is a post office package sorter setting. In this game the player has to pick up packages and put them in the correct sorting box to send them away. Each package has an image on top of it and this image corresponds with a destination. The player has to perform a reach and grasp motion to pick up the package, think about where this package should go and then place it in the right sorting box. By placing the sorting box around the player, the reachable workspace is also measured.

The last idea is to create a virtual puzzle box that the player has to open by solving several puzzles. For example, the player has to touch four hidden buttons to open a latch, and then pull a switch which opens a door to a new puzzle. A good example of such a system is the mobile game The Room (Fireproof Games 2015). The variety of different puzzle actions that can be incorporated in this setting creates a lot of possibilities to use all the functional movements.

Although all three ideas seem a good option it was decided to go with the second one: the post office setting. The first option's omnipresent perspective was considered to be to an unrealistic viewpoint, especially as the game will be played using an HMD, which could demotivate the player. Furthermore, even if there is no real time pressure the fact that there are trains waiting on you removing the objects creates a form of time pressure. The first option also does not have a really strong mind challenging aspect but is more focussed in reaction speed. The third idea does have a strong puzzle aspect to challenge the mind but the setting is unrealistic.

Also the third idea is less supportive of standardized movements which will be necessary for a good, quantitative measurement. The second idea's first person viewpoint on an old-fashioned job combined with the puzzle elements create a much better setting that is realistic, is accessible to different backgrounds and challenges the player's mind. It also uses repetitive functional movements that can be used to make proper measurements. The second option was therefore chosen as the concept to be further developed.

4.1.3. High level concept post office trouble

To further work out the concept, each question of Rollings and Adams (2003, 53) High concept worksheet, discussed in section 2.2.2., is answered. Also, the choices made are reflected back on the requirements shown in Table 7.

What is the nature of the gameplay?

During the game, the player has to pick up items and place them in the right location in a sorting cabinet. An item is displayed in front of the sitting player and he can pick up the item with a normal grabbing motion. The player then has to move the item to a location and release the item there. This creates a reaching and grasping motion that can be used for the assessment of motion disorders. To stimulate the mind of the player, the correct location where the item has to be moved to is given in the form of a puzzle. The sorting cabinet is displayed in front of the player with country names on each box. The item has an image displayed on it and the player has to send the item to the country corresponding with the image (Figure 7). The cabinet is placed in a semi-circle around the player, allowing for the measurement of the reachable workspace of the patient. This gameplay thus incorporates the reaching and grasping motion and measuring at least part of the reachable workspace. Furthermore, when navigating the menu's the player will have to point and reach. The tremor assessment can be conducted when the patient is resting or moving the hands. The range of motion in the fingers can be measured during the actions taken in the game. As required in R5, all the functional motions from Table 6 are thus implemented in the game. The focus is placed on the first two motions in which repetitive movement will help the data retrieval, as required in R2.

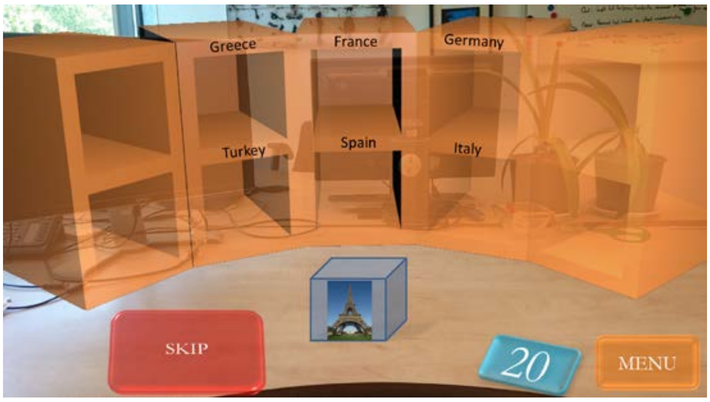


FIGURE 7 - GAME PLAY CONCEPT

The nature of the gameplay was also chosen in this way since it allows the clinicians to manipulate certain parameters so that the patient is pushed to the limit of his motion capabilities, as mentioned in requirement R1. By changing the size and amount of the boxes in the cabinet, the reachable workspace and reaching direction can be more accurately determined. By putting the cabinet further or closer away, the reaching motion can be made easier or more difficult. The size of the package is also an important factor as a very small or very large package can challenge the flexion of the fingers. Therefore it is considered to dynamically adjust the size of the items based on the flexion, range and fluidity in previous results. This can result in packages of a size that the player is unable to move because his motion limitations have been reached.

A skip button has to be provided to allow the player an ingame way of indicating that this item is not possible anymore and even then, the doctor should still have the possibility to change the size according to his wishes. An example on how the size would be changed is shown in Figure 8.

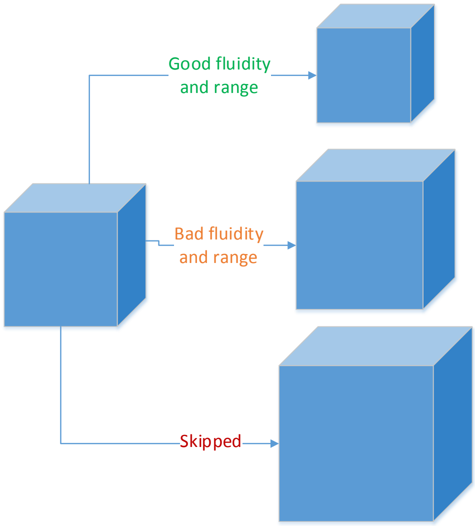


FIGURE 8 - DYNAMIC PACKAGE SIZE DECISION DIAGRAM

The puzzle element on the boxes is an important factor to make the game a challenge to the mind as was required from R7. The images on the items have to be easy to recognise, as this would otherwise hinder the data collection for the assessment since the patient would move less and this is in opposition to requirement R2. The images should therefore be carefully chosen to allow for multiple difficulty levels. It would also be important to create multiple topics. Since each person has different inherent interests (Schell 2005) one person can be very good in identifying cultural attractions while another can recognise the emblems of different soccer clubs. It is therefore important to add these and more topics, allowing the player to choose the ones he likes to play in his session. This will make the game accessible to multiple backgrounds as required in R8.

What is the victory condition of the game, if any?

As it was required that there is no form of time pressure in the game (requirement R10), the goals of the game should also not pressure the player to make haste. Each session of the game will therefore consist out of a number of items that will appear one at a time. The goal of the game is to have as many of the items placed in the correct boxes as possible. The doctor has control over the amount of items in a session as these correlate directly to the amount of measurements made and the length of the session itself. Solving one package is estimated to take between 10 and 30 seconds. This means that a set of 20 packages takes a maximum of 10 minutes. As the game should take no more than 30 minutes (requirement R3) a maximum of three sets of measurements can be performed to gather the data for the assessment, as required by A2. By adjusting the amount of packages or by performing multiple smaller sets of packages the game duration is also easily adjustable, as required from R4.

What is the game's setting?

The game takes place in a post-office on April's Fools day. Someone has encrypted all the international mail with puzzles and they need to be sorted and send to the correct country. This setting was chosen as it is a realistic setting, following the requirement R9, with a slightly absurd twist for entertainment purposes. The game thus takes place in a sorting chamber which consists of a real table and chair plus a virtual sorting cabinet and the virtual packages, which are the items that need to be sorted.

What is the player's role?

The player is the post officer put in charge of sorting the encrypted mail. During the game this officer is constantly busy with sorting and will thus not move away from the table.

What is the player's interaction model?

Current AR technology allows for marker less tracking of hand motions, as discussed in section 2.4. This creates the opportunity to use an interaction model in which physical movements are translated into digital input for the game. A package is thus picked up by making a grasping motion with the fingers and thumb. A button is pressed by touching it with a finger for a second. The input is thus the functional movements from requirement R5. When grasping a package with a motion that is too small or too big, it will not be picked up. If the grasp is correct the package will be linked to the hand and will stay between the fingers while the hand moves to a box. There the package is released by opening the hand. The AR technology therefore creates an interaction model that is natural and intuitive, making it easier for elderly people to get used to the system as was required from R6 while at the same time capturing the data needed for the assessment.

What is the game's primary perspective?

Looking at the examples of AR applications given in the background section, there seems to be two often used perspectives, namely a first person perspective or a fully rotational omnipresent perspective. For the post office trouble game, the first person perspective fits best as the player plays only one role and it is also the most realistic perspective for the proposed interaction model, thus following the requirement in R9.

What is the general structure of the game?

The general structure of the game will consist of five modes:

1. Main menu: from here the player can choose the topics he would like to play (mode 3) and quit the game (mode 5).
2. Tutorial: this mode gives a short demonstration on how the game should be played.
3. Game session: Here the player plays one session of the game.
4. Results: Here the results and correct answers of the previous session are displayed after which the player goes back to the first mode.
5. Game end: After multiple sessions, depending on the clinician, the final results are displayed and the game is stopped

Is the game competitive, cooperative, team-based or single player?

As this design will be the first prototype in an untested system for motion disorder diagnosis, the focus will first lie on developing a single player game. Since a diagnosis is a private affair it should not have other players involved at the same time so it would not be recommended to use the other options in an assessment setting. It should be noted though that if this game is ever adjusted for rehabilitation training it is worth considering implementing cooperative, team-based or competitive modes as the social effects could help improve the rehabilitation.

Does the game have a narrative or story as it goes along?
The story focusses on a post officer who is asked to sort the international mail received on April fool's day. A prankster has removed all the country codes with images corresponding to that country and it is up to the officer to figure out where each package has to be sent. This story was chosen as it combines a realistic job that is widely known with all people with a slightly absurd and funny problem. The goal of this is to foster engagement following the theory of Schell (2005) as discussed in section 2.3. The post officer job was chosen as it is an older profession and has a closer physiological proximity to the target audience. The April fool's joke was chosen as a logical presentation of the puzzle. It was decided to keep the story short and not add any form of story development as this would not fit in a game with a maximum playing time of 30 minutes (requirement R3).

Does the game fall into an existing genre?
The game is a short puzzle game and comes close to the short games commonly played on mobile devices.

Why would anyone play this game?
The game will be played as part of a formal assessment method and patients will thus be asked by clinicians to play the game. The clinicians would want to use this method as it will capture quantitative data in a non-invasive way for motion disorder assessment. At the same time, the game will engage the player with a realistic problem that encourages the mind and memory of the patients playing it, while also inciting them to actively move their body. By using an intuitive interaction model, the game should also create fewer barriers between the game and the target audience. The different topics allow the patient to choose his inherent interest or something in psychological proximity, increasing the engagement value. This will make sure that the player enjoys the experience while the method allows the clinicians to capture the necessary data for their assessment.

4.2. Refinement concept

Before the concept is implemented it was decided to refine the concept further as this would allow us to improve it without spending time on building a working prototype. To do this a new iteration of Meijer's (Meijer 2009) design cycle is started in which the concept is tested and new requirements are derived from these tests. The concept can then be adjusted to account for these new requirements.

4.2.1. Setup, Mock-up, procedure and Participants

To start the second iteration of the design cycle, the current concept had to be tested in some way. It was decided to show a mock-up of the game to a selection of experts who could then give their opinion of the game. Five experts[2] in game design at the TUD where asked to evaluate the concept. In individual sessions, each expert was shown a small presentation about the project, the goal of the game and the mock-up. In the mock-up, each mode of the game was shown using concept images, explaining the full structure of the game. The mock-up can be found in appendix B. After that, the experts were asked if they could review the concept from three angles, namely: Usability, Utility and User experience. The usability of the game concerns if the game is playable by the target audience. The focus here lies on technical aspects that make it harder or impossible to play the game. The utility of the game concerns if the goal of the game is reached and the user experience is akin to the engagement the player has and thus also an important factor. The feedback given was noted down by the researcher for the analysis.

[2] Dr. H.K. Lukosch; Dr. L.J. Kortmann; Ir. A.J. van Veen; Ir. S.A. Tiemersma; Ir. S. Kurapati

4.2.2. Analysis method

The notes gathered from the expert consults were qualitatively analysed. The feedback was sorted per category and then translated into requirements. These and previous requirements were then merged into an updated list of requirements. In this process, each piece of feedback is deemed a possible addition but requirements mentioned by multiple experts should definitely be added to the list. If a new requirement is closely related to another requirement, they were merged into one more detailed requirement. If two requirements are mutually exclusive a choice had to be made. Later tests can then be used to determine the effectiveness of this option, after which the other option can be implemented.

4.2.3. Results

In the discussions with the experts about usability, they found that with a few alterations, the game should be usable. An important issue noted was that when the packages gets smaller, the images on it will also get smaller, making the puzzle unclear and thus unsolvable. In essence, this would make the game a guessing game, something that was not the goal and this issue should thus be resolved. The second issue that was mentioned is that a lot of people will be completely new to this technology. It is thus very important to add a tutorial that not only explains the game but also allows the player to get used to the interactions in the technology.

The utility of the game was difficult to judge without a proper idea on how the measurements will be used exactly to assess the patient. What was mentioned by the experts is that they think it is very important to balance the difficulty of the puzzles. To make good measurements the player should be moving his hands almost constantly which would happen if the difficulty is very easy, but an easy game has no challenge. This would mean that the player loses engagement in the game, making it an irritating experience, and this will most likely influence the motions made and thus the measurements taken. On the other hand, if the game is too difficult the player will not be able to solve any puzzles. Not only will the player loose interest in the game, there will also be no measurements. The difficulty of the game should thus be easy enough for continuous measurement but also difficult enough to keep the player thinking. Since each player is different, this will require some form of difficulty levels in the game that allows each player to play an easy but engaging game. Another utility issue is that the game should measure both the left and the right hand as a motion disorder can affect both or just one. The game should thus ask for each hand separately.

One of the main issues that all experts noticed in the user experience is that the concept needs a bit more gamification. In the concept, the player just sees the amount of packages he still has to do. By adding a scoring system with for example stars or stripes the game will become more engaging. They also said to look into achievements to further stimulate a feeling of satisfaction in the game. In both cases it is important to keep in mind that the players are patients with different levels of motion disorders and that their conditions are most likely going to deteriorate further between assessments. Achievements have to be easy to reach for each player and should thus be chosen carefully so that the motion disorder does not affect it significantly. Another issue is when the player notices decreasing scores compared to a previous assessment sessions. It could have a demotivating effect, especially if it decreases every time. This is also applicable when achievements are coupled to certain records like the smallest package moved. These kinds of demotivating issues should be avoided at all costs.

It should be noted that since they analysed only a mock-up of a game using a novel technology, most experts indicated that it is difficult to fully grasp how things are going to work and look. Thus not every issue in the concept will be found and some issues could in the end be non-existent.

4.2.4. Extended list of requirements

Analysing the feedback from the experts gives seven requirements that can improve the game. The feedback on smaller images becoming too small to see is as issue that is part of a larger requirement, namely that the player should be able to clearly see each image and hear each sound. All visual and auditory aspects of the game should thus be designed so that the player is able to perceive them clearly.

R11: The game has to have clearly perceivable visual and auditory output

The other feedback concerning the players' need to get used to the technology, balanced difficulty and measuring both left and right hand movements are all very good points. Difficulty levels were already a part of the concept to make it accessible for different backgrounds (requirement R8) but tuning the difficulty to optimize movement and engagement is a very important aspect of the design. Therefore it is added as a new requirement. This leads to another three requirements.

R12: The game has to give users time to get used to AR technology
R13: The game's difficulty need to be balanced to optimize continuous movement with maximum engagement
R14: The game needs to incite movement in both the left and the right hand of the player

The last bit of feedback is mostly focused on the scoring system. These points explain the need to find a scoring method that is a bit more gamified and thus helps with the engagement of the players. It needs to fit with the audience and with the previous requirement R9, namely that the scoring is realistic with cartoonish art. As this is a new element to the game a requirement is added to incorporate this in the design process.

R15: The game has to have a gamified scoring system

As mentioned by the experts it is very important that this scoring system is designed in such a way that the player is not demotivated by declining results. As with motor dysfunction, it could very well be the case that a patient sees constant declining motion capability. Demotivation could further lower the results, and thus the measurements, which should be avoided at all cost. This leads to the last new requirement.

R16: The game needs to prevent demotivating situations from happening.

The new requirements are added to the list, shown in Table 8. Again, the conflict in time management, which was already addressed in chapter 3.5, where the short duration of the game may make it difficult to capture enough quantitative data for assessment, is an important issue and this becomes even worse when the player also needs to have time to get used to AR. It is difficult to say if there is truly a conflict here as it is unknown how much data needs to be captured for an assessment and how long this would take. The standard duration of the game might thus have to be adjusted in a later stage to solve this conflict.

TABLE 8 - UPDATED LIST OF REQUIREMENTS

Nr.	Requirement
R1	The game needs to incite players to move at the limit of their motion capabilities
R2	The game needs to be supportive of quantitative data retrieval
R3	The game must not be longer than 30 minutes.
R4	The game's duration needs to be easily adjustable
R5	The game needs to incorporate as many functional movements as possible
R6	The game has to have an easy to use, intuitive interface
R7	The game's exercises need to challenge the mind of the player
R8	The game needs to be accessible for a large variety of backgrounds
R9	The game has to have a realistic setting with a cartoonish art style
R10	The game should have no form of time pressure
R11	The game has to have clearly perceivable visual and auditory output
R12	The game has to give users time to get used to AR technology
R13	The game's difficulty need to be balanced to optimize continuous movement with maximum engagement
R14	The game needs to incite movement in both the left and the right hand of the player
R15	The game has to have a gamified scoring system
R16	The game needs to prevent demotivating situations from happening

4.3. Final concept

To adjust the concept to solve the issues mentioned by the experts, a few alterations where made. First, the player will be able to see the image on the package in full screen when he looks away from the marker used to anchor the virtual environment, which fulfils requirement R11. Secondly, the tutorial is altered so that the player learns to use the new technology and gets time to practice with it (requirement R12). Thirdly, the images are divided into three difficulty levels and the doctor can change these difficulty levels so that the player keeps moving with the optimum difficulty as stated in requirement R13. Fourthly, the packages will be given one of two colours, one for each hand as required in R14. A blue package can only be picked up with the left hand and a purple one with the right hand. This allows both hands to be measured for upper extremity motor dysfunction. Lastly, a more gamified scoring system is added in which the player earns a post stamp for a correct answer (requirement R15). It was chosen to give the player post-stamps as it has a close connection the post-office theme of the game and is still an item collected by many people. Skipped packages are not counted in the score. This leads to the final concept which is shown using a step by step explanation of the game structure following Rollings' and Adams' (2003) procedures. Each step in the game is worked out in specific detail to create an understanding how the game will work. The concept exists out of several modes that represent the various states the game can be in. From each mode, certain actions will lead to moving to another mode until the game ends.

Full game concept:

Mode 1: main menu

The main menu, concept image shown in Figure 9, is the starting screen of the game. The menu exists out of multiple buttons displayed in front of the player. It shows all topics in the game and these can be selected and deselected by holding the hand on the button area for a second. After selecting one or more topics the game can be started by pressing the start button which starts mode 3. Other buttons are the menu button which links to mode M and the tutorial, which links to mode 2.

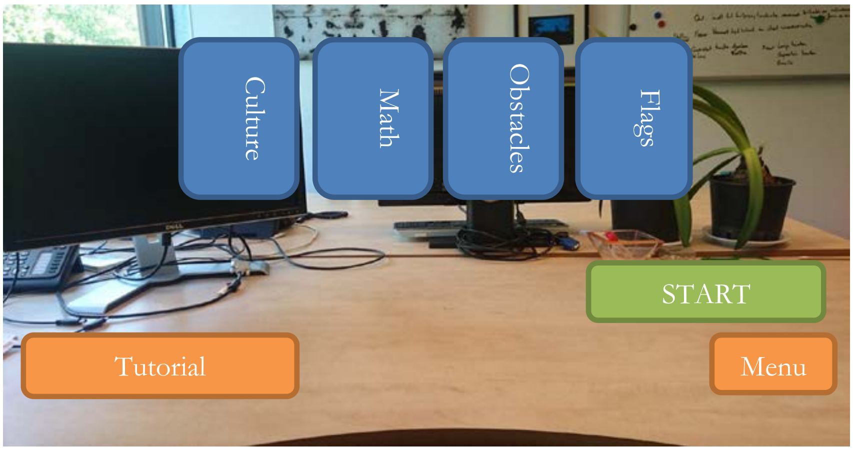


FIGURE 9 - CONCEPT OF MAIN MENU

Mode 2: tutorial

In the tutorial, it is shown how to operate the game. Several actions are explained and displayed using a video of a virtual hand. After each action, the player has to mimic it to enhance the learning process. Actions taken are:

- Move a package to a marked area on the left
- Move a package to a marked area behind the previous one.
- Move the package into a newly displayed box

After the actions are completed the patient can continue to practice until he is satisfied. This is done to give the patient time to get used to the interaction model, as required by R12. During this part, a button with the word "return" is displayed in the right bottom corner which allows the player to return to mode 1.

Mode 3: information and start screen of first set

When a new game is started, the game first shows an information screen (Figure 10). This screen shows the topic chosen, the goal of the game - which is dependent on the topics chosen in mode 1 -, the amount of packages that need to be sorted - which can be changed by the clinician - , a back button to return to mode 1 and a start button to continue to mode 4.

Mode 4: Main game mode

Mode 4 is the main mode of the game. The environment, shown in Figure 7, consists out of the sorting cabinet in a semi-circle around the player at about arm-length distance. The size of the boxes, the amount of boxes and the distance between the cabinet and the player are all variables that can be adjusted by the clinician. This allows the clinician to find the limit of the motion capability of the patient with the game as required by R1. The objects are fettered to a physical marker so that the player can look around in the virtual environment. Other elements visible are a counter counting the packages still to send, a score area showing the amount of post stamps collected and a menu button. These are all fixed to the same position in the screen, namely the bottom right, instead of being fettered to the real world. The last thing visible is a skip area fettered to the real world, on the table in front of the cabinet.

FIGURE 10 - CONCEPT OF MODE 3

When the game starts, a package appears in front of the player. Country names appear on the sorting boxes and, depending on the chosen topic, an image is shown on the package indicating which country it needs to be send to. The player can pick up the package by making a grabbing movement big enough to encompass the package. The colour of the package indicates if it has to be picked up with the right hand, a purple package, or the left hand, a blue package, as required in R14. When the package is grabbed, the package is linked to the hand and the player can move it around. The player then moves the package to one of the boxes and the package is transferred to the box, after which it is moved backwards and is removed. If the package is send to the correct location, a post-stamp is given to the player. The size of the package is altered dynamically based on the success of previous packages to help identify the motion limits (R1) or the size can be altered by the clinician. If the current package cannot be picked up, the player can put his hand down in the skip area and a new easier package is created without penalty, as required by R16. If the player has trouble seeing the image on the package, he can look away from the marker to blow up the image to the full size of the screen, a required by R11.

During play, the player can pause the game by hovering over the menu button to activate mode M. After all the packages have been sorted the game will automatically go to mode 5.

Mode 5: Game results
Mode 5 is another pop-up screen, shown in Figure 11, that displays the amount of post-stamps gathered and a list of each image, the correct answer and the answer given. A continue button sends the player back to mode 1.

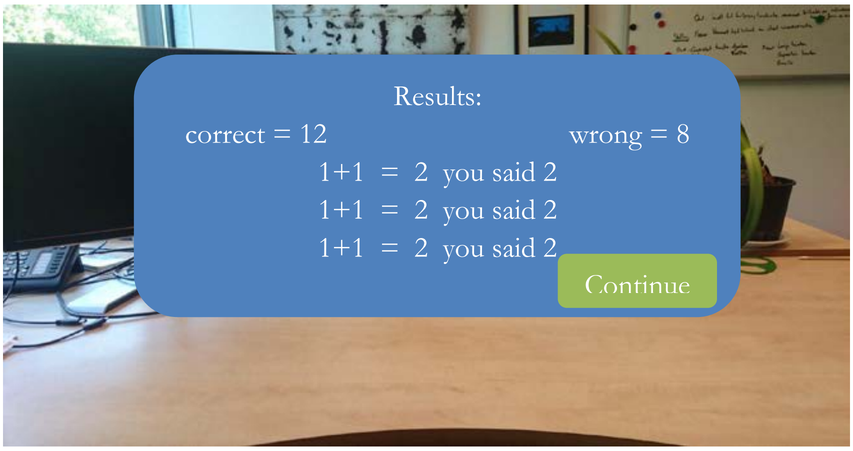


FIGURE 11 - CONCEPT OF MODE 5

Mode M: Menu

Mode M is a list of two options. Continue game sends the player back to the previous mode. The second option depends on the mode from which the menu is accessed. If in mode 4, the quit button sends the player to mode 5 while only registering the finished part of the packages. If in mode 1, the End session button will end the game and send the player to mode end. The clinician can at any time quit a session from his interface.

Mode end: final results

Mode end shows the player what he has achieved till now. It displays a list of game sets played in current and previous sessions with the activated topics and the total amount of stamps gathered. The clinician can then shut down the game from the clinician interface.

5. Usability and engagement study

Now that the game concept is developed, it has to be found out if the design truly facilitates engagement. To do this, a prototype of the game is developed and then tested. The goal of this experiment is not only to understand if the game is engaging but also to create data that can be used in the second iteration of Meijer's (2009) design cycle to find new requirements. Therefore, multiple tools are used to test different aspects of the game. This chapter will first explain how the concept was implemented in a prototype, then it discusses the participants for the tests and continues with the procedures and tools used. Then the results of the different tools will be shown and lastly these results will be discussed to find out if the goals of the experiment have been reached.

5.1. Game implementation

The first step of testing the game is implementing the concept into a working prototype. During the programming of the game there were some troubles with optimizing the hand recognition. As this is of great importance for the performance, the hand tracking was prioritised over the secondary modes of the game. Therefore, the implemented game only consists of the main phase of the concept and all the surrounding menus, topic choices, tutorials and result pages were excluded. In the main phase the player has to move ten packages of five by five by five centimetres into one of ten boxes, shown in Figure 12. Each box is labelled with a country name and each package with a corresponding image. Two puzzles were available: one with different cultural attractions and one with different soccer club logos which are shown in appendix C. As there were only three pictures per country available no difficulty levels were implemented but the images where chosen so that each country has at least an easy and a difficult image. Lastly, the game was only capable of recognising the right hand and the colour identification for different hands was therefore not implemented. After ten boxes were sorted, a message was shown with the amount of packages that where correctly delivered and the amount wrongly delivered which is also the final screen of the current implementation.

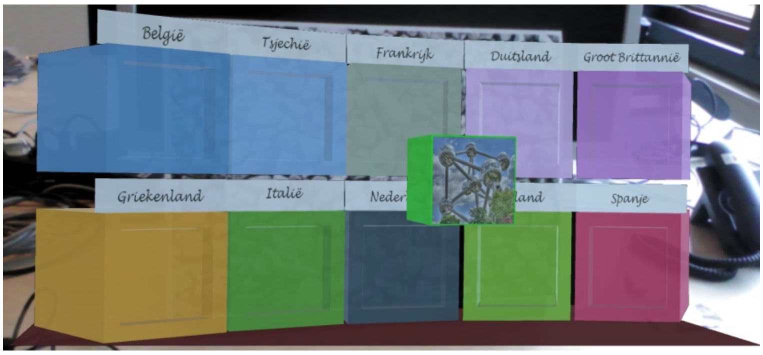


FIGURE 12 - GAMEPLAY VIEW

5.2. Measures

To test if the game is truly engaging, two validated questionnaires, one on usability and one on engagement are used as measurements (Brooke 1996; Ijsselsteijn et al. 2008). Next to these two, ten personal questions about age, career and computer experience were also included. These questions are shown in Table 9.

Table 9 - Personal questions (QP#)

#	Question	Answer options
1	What is your age?	Open
2	What is your gender?	Open
3	Are you left or right handed?	Left; Right
4	What is, or was, your occupation?	Open
5	What is your experience with Computers?	Scale from 1 to 5
6	How many times have you used AR technology	Never; 1; 2-5; 5-10; >10; I do not know
7	Are you quick to adapt to the use of new technologies?	Scale from 1 to 5
8	Do you play digital games?	Never; singular; monthly; weekly; daily
9	On what type of platform do you play these games?	Mobile device; Computer; Wii; Playstation or Xbox; other
10	Do you play analog games?	Never; sometimes; monthly; weekly; daily

The usability questionnaire used is the System Usability Scale (SUS)(Brooke 1996). SUS was chosen because it gives a validated score without subjecting the interviewee to a lot of questions. As this research is meant to measure engagement we wanted to keep the usability questionnaire small. As SUS only has ten questions it was well suited in this regard. The SUS questions are shown in Table 10. All questions are statements that are answered on a Likert scale of 1 to 5 with 1 being “completely disagree” and 5 being “completely agree”. With the SUS, odd questions are formulated positively and even questions are formulated negatively so to not create a bias in the answers (Brooke 1996).

Table 10 - - SUS Questions (QS#)

#	Question
1	I think that I would like to use this system frequently
2	I found the system unnecessarily complex
3	I thought the system was easy to use
4	I think that I would need the support of a technical person to be able to use this system
5	I found the various functions in this system were well integrated
6	I thought there was too much inconsistency in this system
7	I would imagine that most people would learn to use this system very quickly
8	I found the system very cumbersome to use
9	I felt very confident using the system
10	I needed to learn a lot of things before I could get going with this system

To measure the engagement of the participants, the Game Experience Questionnaire (GEQ) was used (Ijsselsteijn et al. 2008). IJsselsteijn et al. (2008) developed this questionnaire to measure the experience and actually never uses the term engagement. Still, the seven dimensions of player experience used by Ijsselsteijn, described in Table 3, are elements that match with being engaged, as is discussed in section 2.3. Especially the Flow dimension is important as it corresponds with keeping the interest of the player which is important in engagement. GEQ scores should thus reflect the state of engagement of the player. Although sources state that this questionnaire is validated no actual paper about this validation has been published (Ijsselsteijn et al. 2008; Norman 2013). Unfortunately, there are not many other engagement questionnaires out there. Kent Norman (2013) describes a different, validated survey called the game engagement questionnaire but as this one is designed for detecting people prone to being negatively influenced by games there is a certain bias in the questions. This bias makes this questionnaire not appropriate for this research and therefore it was chosen to use the GEQ. The GEQ exists out of multiple modules, namely a core module, a short in-game module, a social presence module and a post-game module. As the focus of this research is to study the engagement of the player during gameplay, only the 33 questions of the core module where used. The in-game module was not used because it is most useful when comparing different gameplays or modes, requiring short questionnaires in between the parts, and this was not the case. The core-module questions are shown in Table 11. Each question is a statement and the participant answers how much this statement was the case on a Likert scale from 1 to 5 with 1 being "not at all" and 5 being "extremely". As all the participants where Dutch citizens and it could not be assumed that they all have a sufficient level of English available, an official Dutch translation of the questions was used. These translated questions are shown in appendix D.

TABLE 11 - GEQ QUESTIONS (QG#)

#	Question	#	Question
1	I felt content	**18**	I felt imaginative
2	I felt skillful	**19**	I felt that I could explore things
3	I was interested in the game's story	**20**	I enjoyed it
4	I thought it was fun	**21**	I was fast at reaching the game's targets
5	I was fully occupied with the game	**22**	I felt annoyed
6	I felt happy	**23**	I felt pressured
7	It gave me a bad mood	**24**	I felt irritable
8	I thought about other things	**25**	I lost track of time
9	I found it tiresome	**26**	I felt challenged
10	I felt competent	**27**	I found it impressive
11	I thought it was hard	**28**	I was deeply concentrated in the game
12	It was aesthetically pleasing	**29**	I felt frustrated
13	I forgot everything around me	**30**	It felt like a rich experience
14	I felt good	**31**	I lost connection with the outside world
15	I was good at it	**32**	I felt time pressure
16	I felt bored	**33**	I had to put a lot of effort into it
17	I felt successful		

Next to the questionnaires, two qualitative methods were also applied to gather feedback for the next iteration of the design cycle. First, the test sessions were recorded to capture actions taken and the expressions made by the participants. This is used to identify if the patient is agitated and engagement is failing, but also to identify where there are issues with the usability, identifying problems to solve in the following iterations. Secondly, the participants were also debriefed to discuss problems they experienced during the test session. Thus, in an unstructured interview, the patient is asked his opinion on the game and was asked to elaborate on observations made during the session.

The different measurements taken result in four different data streams: the SUS scores, the GEQ scores, the observations and the feedback from the debriefing. The results of this data is collected and analysed to first find any hint of the engagement the players felt during play. For this part, the quantitative data of the questionnaires will be especially useful. Secondly, the data is analysed to find new requirements for the next iteration of the game design. Here, the qualitative data from the observations and the debriefing will be most helpful.

The quantitative data from the questionnaires is treated following the validated method which is part of the tool. For SUS, the separate questions do not have much power therefore an average SUS score is calculated. First all individual scores are transposed to a 0 to 4 scale. Therefore for the odd questions the score minus one is taken and for the even questions five minus the scores is taken. The sum of these scores times 2.5 gives a usability ranking between 0 and 100 in which a score over 67 is assumed as a positive score (Brooke 1996).

TABLE 12 – QUESTION NUMBERS CORRESPONDING WITH ASPECTS

Aspect	Question numbers
Competence	2, 10, 15, 17, and 21
Sensory and Imaginative Immersion	3, 12, 18, 19, 27, and 30
Flow	5, 13, 25, 28, and 31
Challenge	11, 23, 26, 32, and 33
Positive Affect	1, 4, 6, 14, and 20
Negative Affect	7, 8, 9, and 16
Tension/Annoyance	22, 24, and 29

For the GEQ, seven different aspects of experience can be calculated by taken the average of certain questions, which are shown in Table 12. These aspects are used to better explain how the game is experienced and thus how engaging a game is. Competence shows how competent a person feels in the game. Sensory and imaginative immersion is about how the person found the game world and how much he felt that he was in it. Flow is about how much the player is drawn into the game, forgetting everything else around him. Challenge shows how much the player felt positively challenged to complete the game. Positive affect is about how much positive emotions the game aroused in the player whereas negative effect is the amount of negative emotions aroused by the game. Tension and annoyance is, of course, the amount of irritation the player felt during play, which is closely related to negative affect but shows a distinct difference in that tension is actively aimed at the game whereas negative affect is a passive feeling, like boredom. The scores of the seven aspects are compared between players to understand the average engagement the players felt. For a positive score, the first five should have a high rank and the latter two should have a low ranking.

The qualitative data is analysed by first watching the recording and noting down anything that happened during the game and any feedback mentioned during the debriefing. The written logs of the different participants are then compared to find common points and these points are then transformed into requirements and positive feedback. These are then ordered based on the amount of participants mentioning this requirement. Only the requirements used by multiple participants are used for further development of the design. This prevents an opinion of one person to affect the concept. It could be that only one person finds a flaw it the design but this should then be noticeable in later iterations of the game.

5.3. Setup

The first step in building the test setup was to decide which hardware was best fitted for the job. The most important decision here was to choose which off-the-shelve sensor would be used. There are many different sensors on the market and they can vary in the method used to track hands and the space dimensions in which the tracking is optimal. For the purpose of the game and the TIM project, it is important that the sensor is as accurate as possible and can track the different joints in the hand for proper measurements. After some research and testing it was decided that the Intel® RealSense F200 RGB-D Camera was the best option currently available as it can track 21 points on each hand with the smoothest life tracking (Intel.com 2015). The separate sensor was joined with the Epson Moverio BT-200 OST-HMD to create the necessary hardware for the game (Figure 13). The game was implemented in Unity3D and ran on a Dell Precision M4800 Laptop in the Editor Mode of Unity3D. Using the Intel® RealSense API for hand tracking and the Vuforia Target Tracking API for marker detection, the game could then be played on the HMD through the Unity 3D android remote play app. A marker is placed in front of the player so that the Vuforia API can fetter the virtual environment to a real location (Cidota et al. 2015).

FIGURE 13 - THE EPSON HMD JOINED WITH THE INTEL SENSOR

During a test session the participant was seated at a desk with the marker in front of him. The participant wears the HMD which is connected to a laptop. The laptop is also on the table so that there is enough cable length for free movement. The marker is slightly raised so that it is straight in front of the participant, making sure he does not have to bow done to see the AR environment. A Canon legria FS36 camera was set up at the side of the table so that it could capture both facial expressions and the arm motions made. The setup is shown in Figure 14.

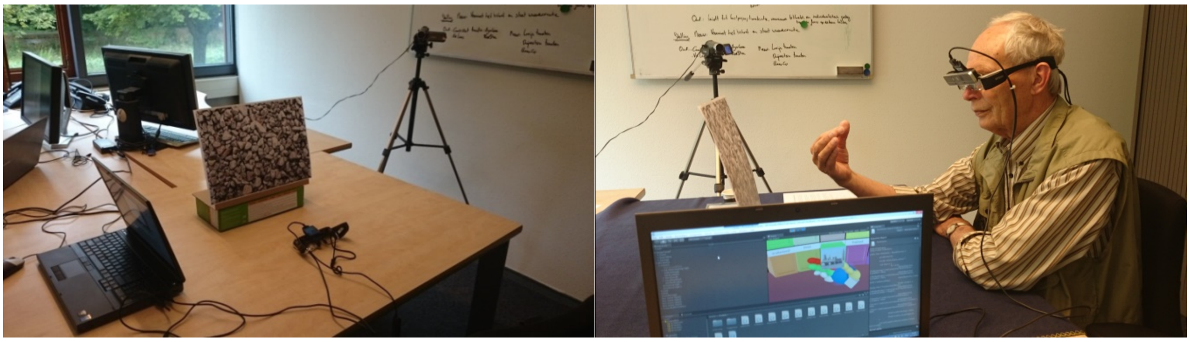


FIGURE 14 - TEST SETUP

5.4. Procedures

The following procedure was used during the experiments. First, an explanation of the TIM project, the research and the goals of the experiment was given. Then the participant was asked if he or she was okay with being recorded and, if yes, two informed consent forms were signed. The participant then fills in the personal information part of the questionnaire while the researcher starts the recording. The researcher then shortly explained how the game worked and what was expected of the player. The participant was given a bit of time to get used to the technology. The participant then plays the prototype of the game discussed in chapter 5.1. After playing the game the participant is asked to fill in the SUS and the GEQ respectively. Finally, the participant is debriefed, the recording is stopped and the participant is given one of the informed consent forms before he or she leaves.

5.5. Participants

To perform the tests a sample group was needed that closely resembles the target audience. The tests could not be performed on actual patients as this would require a long request procedure at the ethical commission of the LUMC, which is too time consuming for the scope of this project. Also, as it was still unsure how well the tracking of the setup would perform it was decided that it was better to perform the tests with healthy people, focusing more on the mental aspects of the game. Still, the sample group should resemble the patients as much as possible and therefore people older than 50 years and with a large variety in backgrounds were chosen. During the experiments, eight participants between 52 and 79 played the game. Five of these participants are below 65. Half of the participants where female and there was a large diversity in occupations which reflects the variety in backgrounds.

5.6. Results

In this section the results of the questionnaires, the observations and the debriefing will be shown. The raw data can be viewed in appendix D. The personal data given showed that the computer knowledge of the participants was diverse as both their pc experience and their affinity to technology have a large variation in answers. However, none of the participants have used an AR system before and only half of them indicated to play games and then only short mobile games.

5.6.1. SUS

The calculated SUS rankings show a large variation in results. The usability indicated by the participants, ranges from 28 to 90 with a mean of 60 and a standard deviation of 21. When looking at the boxplots of the individual questions, shown in Figure 15, the same spread in results can be seen in the high standard deviation. Of note is that the odd, positive formulated questions are all around average with QS1 a bit lower and QS7 higher, whereas the even, negatively formulated questions are all ranked below average with three of them going significantly lower.

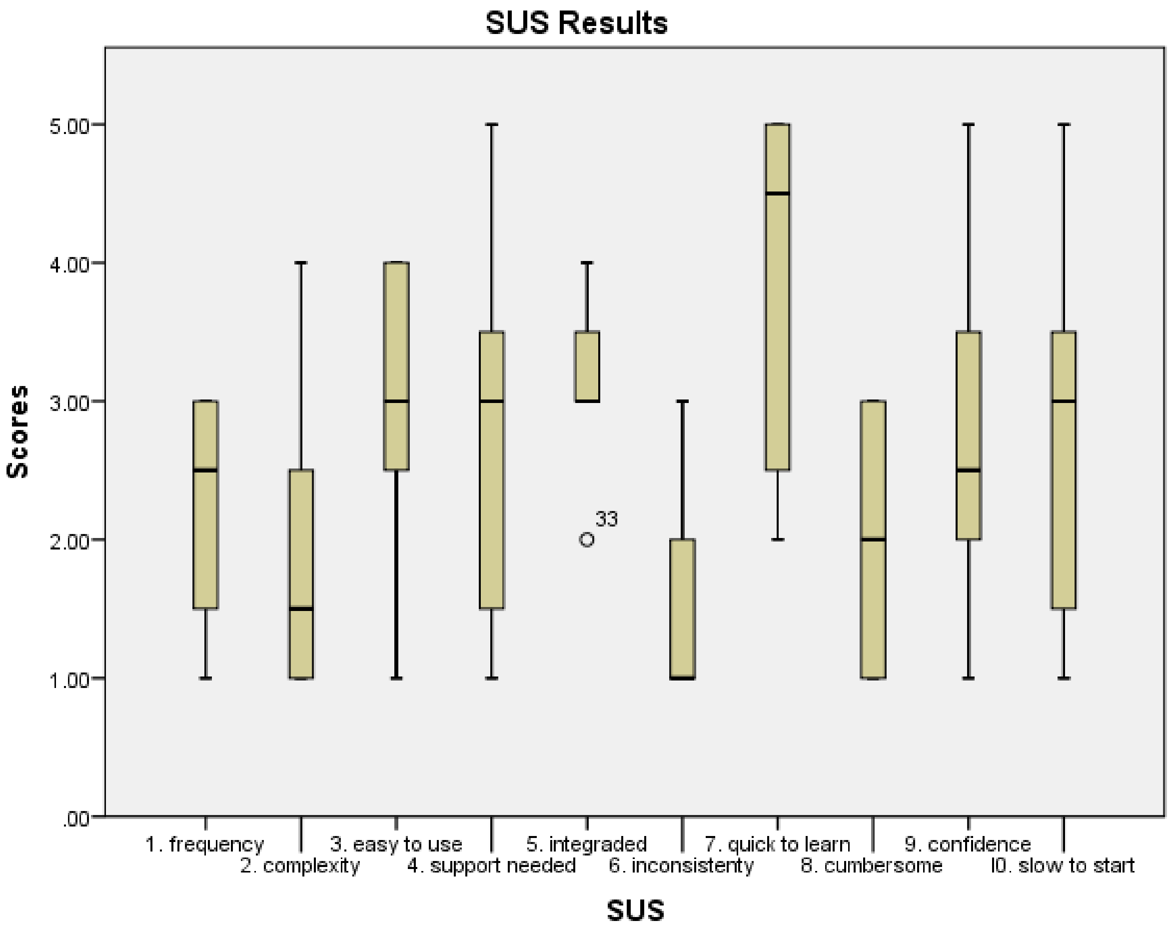


FIGURE 15 - SUS RESULTS (N=8)

5.6.2. GEQ

The GEQ gives scores in seven categories. The boxplots of these categories are displayed in Figure 16. The first five categories require a high score but on average, all scored between 2 and 3. Competence and immersion do have a larger deviation but the majority of scores was below 3. The challenge factor was rated by all participants to be below 3 and positive affect was mostly rated on three with a few below. Flow had a better result as it was rated between 2 and 4 with an average slightly above 3. The two categories requiring low ratings did have low results, with all participants saying they did not feel negative affect. Tension was rated a bit higher between 1 and 2 but the majority off the participants rated tension with a 1 as well.

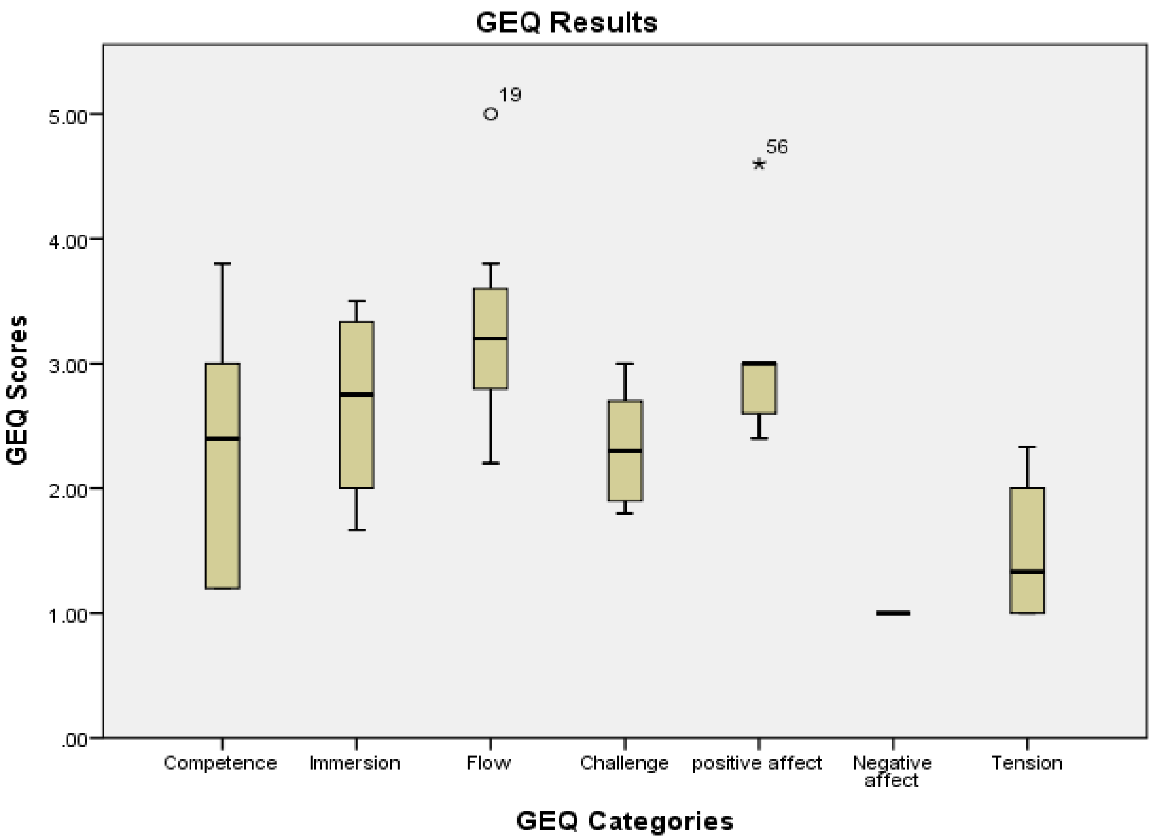


FIGURE 16 - GEQ RESULTS (N=8)

5.6.3. Observations during test sessions

The patients where observed to see how they reacted to their progress in the game and if there are any specific problems they encounter. In general, each participant showed willingness to play the game and started with getting used to the technology. This start was usually with a bit of irritation which lasted for a bit, depending on how quick the patient could get used to the technology. After that, patients stayed focused on the package and as it got easier, people showed positive affect. It was also observed that multiple participants held the HMD up with their left hand. Another thing is that two participants showed that they had difficulty with the puzzles, expressing that it was not a subject they knew much about. Two others found the puzzles easy and were finished quite quickly. Three participants did express some surprise about the amount of knowledge they did possess. Three of the participants had trouble with perceiving the depth in the virtual environment, often grabbing in front or behind the package or stopping in front of the sorting boxes, thinking they were already there. Participants also had trouble getting their hand recognised. During the session they had to get used to which positions are recognised and which are not until in the end of the session things worked better. One person managed to move all packages with the back of the hand towards the sensor where for everybody else this would have resulted in loss of tracking. Another issue found was that people had trouble coordinating the hand with the head. As the sensors are placed on the HMD, the player has to keep the hand in his sight so as not to lose the package. In real life though, people tend to look to the goal of the objective so when the package has been picked up people often focus on the correct box, losing the package out of sensor sight. The participants got used to it with some practice except for three of them, who had trouble adapting and often lost the package and hand out of sight.

Two of the participants also had some initial trouble with enlarging the image by looking to the side. Their intuitive reaction was to bring the package closer which often resulted in obscuring the marker from the camera, thus losing the virtual environment. One of them got used to looking to the side and the other did not. Three participants who had glasses tried to wear the HMD over it but this did not work and the participants choose to do it without their own glasses.

5.6.4. Debriefing

Several issues were raised by multiple people and in Table 13 each of them is shown, ordered on the amount of people expressing that issue. To complete the table, issues found during the observations are also added. The first issue raised is that the sensors work best as long as they see the palm of your hand. Four participants expressed that they found this movement very difficult and unnatural, especially when putting the package away in the boxes. Another raised issue by three of the participants was that the texts and images were not recognizable enough. For the text on the labels this means that it was not big enough and the font was also not recognisable enough. The problem with the images was already foreseen but the solution with looking away to see the bigger picture was not working for them. One participant stated that looking away to see the big picture felt like stepping out of the game and then stepping back in again, so he avoided using this solution. Three of the participants also found the AR glasses cumbersome as they were heavy and had a tendency to slip on the nose which then misaligned the display with the eyes. Also mentioned by three of the participants was the need for more practice before playing the game. The main reason for this was to get more time to get used to the technology. The last issue was that four people felt that the game was confined to a small space because the boxes where displayed in a square in front of them resembling a screen. This limitation caused participants to keep their head focussed on the middle, sometimes even dropping the hand out of sensor range when trying to reach a box. It should be noted that three of the participants also clearly stated that the found the game a fun experience.

TABLE 13 - - LIST OF PARTICIPANT ISSUES

	Issue	# people
1	Unnatural movement and grip	4
2	Feeling confined to a screen	4
3	Image and text not visible enough	3
4	Heavy AR HMD	3
5	More practice would be helpful	3
6	Difficult to perceive depth	3
7	Difficult head-hand coordination	3
8	Not glasses friendly	3

5.7. Discussion

The goal of the experiment described in chapter 5 was twofold. The first goal was to determine if the game is engaging, the second goal was to create a new set of requirements to update the concept with. To reach these goals, the concept of the game was implemented in an AR system and tested by eight healthy people older than 50. To measure the engagement two questionnaires, the SUS and the GEQ, were used. To improve the concept, the participants were recorded during their play session and they were debriefed afterwards. The SUS results show a usability of 60 on a scale to 100 with a large standard deviation of 21. The GEQ scores show that the game is has an average score in the positive attributes and a very low score in the negative attributes. The observations and debriefing of the play-sessions show that the participants did like to play the game and also resulted in a list of 8 issues (Table 13) that could be improved. The two goals are discussed separately in the following sub-chapters.

5.7.1. Engagement of the game

IJsselsteijn et al. (2007) mentioned that a good game design should always take two factors into account: usability and user experience. A good game needs both of these factors and although a game can only be usable, it cannot have a good user experience without being usable Therefore both the SUS scores and GEQ scores need to be taken into account when discussing the engagement of the game.

Brooke (1996) indicates that for a product to be usable the SUS ranking should be above 67 on a scale to 100. The current results do have some results above but with an average of 60 and a standard deviation of 21 most of the results are negative. There is thus still a lot to improve before the game can be called usable but the question remains why the SUS results are so wide spread. Plotting the usability scores against the PC experience of the participants, as is shown in Figure 17, shows that there could be a positive link between SUS and computer experience but because of the low sample size and large spread, a larger study needs to be performed before any real conclusions can be drawn. Another way to explain the large variation is because all the participants had never played with an AR application before. This would mean that they did not have any comparative material and their usability scores are thus based on the expectations the participants had. As AR technology is a novelty to them, these expectations will be different with each participant and their answers will thus be more spread than if they would rate a normal computer application. This could explain the spread but does not diminish the fact that the usability is still low and needs further improvement.

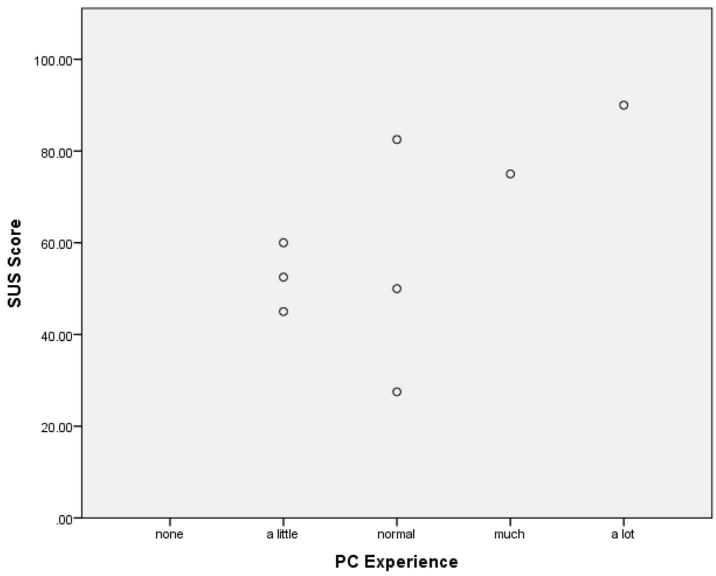


FIGURE 17 - SUS RESULTS VERSUS PC EXPERIENCE

The GEQ results show less spread but the results are not encouraging. They show that the participants felt the positive aspects somewhere between “slightly” to “moderately”, indicating that they were not that engaged by the game. At the same time, they did not feel negative affect nor was there any tension created by the game. If the game was simply bad, it would be expected that these negative aspects would score higher. This makes it difficult to understand what the performance of the game was. Three possible scenarios could explain the results gathered from the experiment.

First, there is the possibility that the GEQ results are not influenced by any other factors. The game is simply interesting enough to not get bored or annoyed but is also not really that engaging. This would mean that the focus of further research should be on improving the engagement of the game. The second option is that the game is not engaging at all and that the current scores are a result of the participants' interest in the novelty of the AR technology. In this case the best course of action would be to design a new game that is more engaging. The last option is that the game is engaging, but that the lower usability is suppressing the positive aspects of the experience for the participants. The game would then have potential for engagement but further research should focus on upgrading the usability of the game. Unfortunately, with the current sample size and experiment results, it is impossible to say which of the options is correct.

5.7.2. Updating requirements

By analysing the issues of the participants from Table 13 and the current requirements from Table 8, a final list of requirements can be made. Two of the issues, point 4 and 8, deal with hardware design, indicating that the participants found the HMD too heavy and found it bothersome that it was not possible to wear their own glasses below the HMD. Since the goal of this research is the game design, issues with the hardware design are not part of the research but are of course important. Therefore these are kept separate as recommendations for the TIM project.

Two other issues that are more focused on the technology used are issue 1 and 6. The unnatural movements that have to be made are a consequence of the sensors and tracking algorithms. The sensors track 21 points in the hand and match them to a model of the hand. This method is calibrated to work best when the palm of the hand is turned towards the sensor, which results in the player having to make the unnatural movements. To resolve this issue, more time needs to be invested in researching the tracking method and developing it to make it also recognise the back of the hand.

R17: The game's tracking algorithms needs to be optimized for both the front and the back of the hand

Issue 6 was that the participants had difficulty to perceive depth. The HMD also has the capability to create a 3D virtual environment which would be perfect to enhance the depth perception of the players. The concept already assumed that a 3D environment would be used but this was not part of the requirements. The participants having trouble perceiving depth only strengthens the importance of 3D and it is therefore added to the list.

R18: The game needs to be displayed in full 3D

Issue 3 in Table 13 indicates that the participants found the images and text not visible enough. Requirement R11 in Table 8 already deals with this issue and a solution to this was implemented. The current solution to this requirement, looking away from the marker to see a blown up image, is thus not good enough. The labels should be made bigger, not be cursive and a better readable font should be chosen. A better way to enlarge the pictures within the game setting should also be found. This is deemed an important requirement for the next iteration and therefore is placed separate of requirement R11.

R19: The game has to have an intuitive way to enlarge the image.

Issue 5 indicates that the participants wanted to have a bit more practice before starting the game. As the original concept includes a tutorial which could not be implemented in the prototype this issue was already solved in the concept. It does strengthen the importance of using a slow tutorial which gives ample time for new players to get used to AR technology. The trouble with head-hand coordination experienced by the participants, issue 7, could also be solved with some extra practice.

Issue 2 deals with the player feeling that their vision is confined to a certain area, which made them afraid to look away from the marker. This feeling is most likely due to the players being accustomed to normal screens and the virtual world being about the size of the screen. The original concept already included the idea of placing the boxes in a semicircle and this issue only strengthens the importance of this. If the virtual environment encompasses around the player, he is forced to look around which solves the issue. This will also make the movements made less unidirectional and will improve the measurements made. The encompassing virtual environment is therefore a new requirement for the game. This requirement does bring a technical problem into the fold. To create a virtual environment that augments reality, it needs something real to fetter it to. In the current prototype a marker was used but this one will be lost from sensor range when looking to the side. So next to the requirement for the concept it is also recommended that a way to create an encompassing 3D environment is researched.

R20: The game has to have an encompassing virtual environment with a semicircle sorting box

By adding the new requirements to the old ones the final list of requirements is made which is shown in Table 14.

TABLE 14 - FINAL LIST OF REQUIREMENTS

Nr.	Requirement
R1	The game needs to incite players to move at the limit of his motion capabilities
R2	The game needs to be supportive of quantitative data retrieval
R3	The game must not be longer than 30 minutes.
R4	The game's duration needs to be easily adjustable
R5	The game needs to incorporate as many functional movements as possible
R6	The game has to have an easy to use, intuitive interface
R7	The game's exercises need to challenge the mind of the player
R8	The game needs to be accessible for a large variety of backgrounds
R9	The game has to have a realistic setting with a cartoonish art style
R10	The game should have no form of time pressure
R11	The game has to have clearly perceivable visual and auditory output
R12	The game has to give users time to get used to AR technology
R13	The game's difficulty need to be balanced to optimize continuous movement with maximum engagement
R14	The game needs to incite movement in both the left and the right hand of the player
R15	The game has to have a gamified scoring system
R16	The game needs to prevent demotivating situations from happening
R17	The game's tracking algorithms needs to be optimized for both the front and the back of the hand
R18	The game needs to be displayed in full 3D
R19	The game has to have an intuitive way to enlarge the image
R20	The game has to have an encompassing virtual environment with a semicircle sorting box

5.8. Learnings and Limitations

During the thesis research, possible solutions for the requirements in Table 14 were identified. In some cases, problems were encountered and other solutions had to be found. This section describes what was learned from each requirement and also discusses the limitations of this study. As requirement R17 to R20 were identified in the previous section, it will only discuss possible ways to implement them in the game.

R1: The game needs to incite players to move at the limit of their motion capabilities.
This requirement forms the basis of the assessment capabilities of the game. To conduct the assessment, the patient needs to show what he can do. This requires that the game asks the full possible range of each motion. The word incite is used because the game should not force the action but instead needs to set the patient up to make him try to reach the limit of his capabilities. In the Post Office Trouble game this is done with two elements, discussed in chapter 4.1. The first element is that the packages that the patient needs to pick up are dynamically altered in size based on previous results. This allows the game to identify the limits of his grasping motions. The second element is that the clinicians can alter the parameters in the game at any given time. If the patient is having an easy or hard time, the clinician can alter the parameters to actively search the limit in motion capability.

R2: The game needs to be supportive of quantitative data retrieval.
To gather quantitative data will require multiple measurements. Furthermore, as is discussed in section 2.4, hand tracking technology is still in development. This means that this technology is still somewhat inaccurate and repetitive motions are needed to measure accurate data. The exercises in the game are chosen so that the player is performing repetitive movements (section 4.1). As this thesis did not yet test the data retrieval nothing more can be said about its effectiveness.

R3: The game must not be longer than 30 minutes. And **R4: The game's duration needs to be easily adjustable.**
To keep the duration of the game short and adjustable, it uses several game sessions, as discussed in section 4.1. The package sorting setting allows this type of game as each session consists out of sorting a set of packages. Each session will take about 10 minutes and the clinician can than string together as many sessions as he deems necessary.

R5: The game needs to incorporate as many functional movements as possible.
In section 3.2, five functional movements were identified that need to be measured for a motion disorder assessment. It was decided to focus the exercises of the game on two of these movements (section 4.1). Reaching and grasping is performed when the player picks up a package and places it in a box. The reachable workspace is measured when the player has to reach several different boxes. During these exercises, measurements of the joints in the hand can be used to measure two other functional movements: the range of motion in the finger joints and the tremor in the hands. The last functional movement, reaching and pointing, is performed when navigating through the menu and screens. This does not happen much so this functional movement is not measured as prominently as the others. One of the main issues found in chapter 5.6 is that the current implementation results in an unnatural hand position. This position makes it difficult to properly perform the functional movements which should be improved before the hand recognition the game is useful for motion disorder assessment.

R6: The game has to have an easy to use, intuitive interface.
To make the interface as intuitive as possible, the actions taken are modelled as much as possible to normal, physical actions (section 4.1). Grasping and moving the box is done with grasp and move actions, buttons in the game are activated by touch and to look around is done by head movement. The observations made in chapter 5.6 show that although it was the first time the participants played the game, they understood the interface quickly and intuitively.

R7: The game's exercises need to challenge the mind of the player.
To engage the players, a puzzle element is added to the exercises. The interviews from section 3.3 show that the target audience likes to be challenged with brain exercises. Therefore, the packages are all given an image that corresponds with a country. The player then has to send the package to the corresponding country to successfully complete the puzzle. This type of puzzle was chosen to practice memory and general knowledge of the player.

R8: The game needs to be accessible for a large variety of backgrounds.
As people with motion disorder can come from any kind of background, the puzzle should also be playable by all people. To allow for this there are multiple puzzle topics available, as explained in section 4.1. The player can choose multiple topics that fits his interests and these will be used in the puzzle. Two examples of topics are shown in appendix C.

R9: The game has to have a realistic setting with a cartoonish art style.
To keep the game realistic, the post office sorter setting was chosen. The sitting post officer corresponds to the sitting position of the player. The HMD allows the player to see exactly what the post officer would see and the actions taken are also as realistic as possible. The cartoonish art style was planned but was not implemented in the first prototype. Therefore this was not further mentioned in this report but should still be implemented in later prototypes.

R10: The game should have no form of time pressure.
To remove all forms of time pressure, the packages are given one by one and no timer will be visible. The goal of the game is focussed on sending the packages to the right location. No focus is placed on the time it took to send them. The danger here lies in the conflict between this requirement and requirement R2 and R3. The game needs to take enough measurements in the 30 minute time span and at the same time the player should not be pressured to hurry up. For the clinicians, it is thus very important to keep the player at ease as this is the most beneficial for his engagement.

R11: The game has to have clearly perceivable visual and auditory output.
Based on the feedback received in section 4.2, clearer visual and auditory output was required from the game. For now, there is no auditory output but if this is added later it should be taken into account that as elderly people often have lesser hearing, the sound should be loud and clear. For the visuals, the text and images in the game need to be large enough to be readable. As the packages can become quite small it was decided to add a function where looking away from the marker would show a blown-up view of the image on the package. The feedback given in section 5.6 told two things. Firstly, the text was not clear. It should be made bigger and the font was also difficult to read. This shows that using clearly perceivable letters is just as important as the size of the letters. Secondly, the solution chosen for enlarging the image wasn't used that often. The intuitive response is to draw the package closer so using this in the interaction model will be a better solution.

R12: The game has to give users time to get used to AR technology.
As AR is a new technology, the game experts (section 4.2) expressed the importance of getting used to it. The concept therefore included a tutorial mode in which the player was given step for step instruction on moving a package. It also gave the player time to practice these movements. Due to limited time this tutorial was not added to the implementation (section 5.1). Instead, a participant was given a few instructions before the test started (section 5.4). The responses in the debriefing in section 5.6 reinforce the importance of requirement R12 as the participants would have liked to have more time to practice before playing the game. A tutorial is thus vital and it has to give the player time to practice.

R13: The game's difficulty need to be balanced to optimize continuous movement with maximum engagement.
AS discussed in section 4.2, the game needs to incite as much movement as possible while keeping the player engaged. This requires the difficulty of the game to be balanced in such a way that it is easy enough for the player to solve the puzzle in a second and hard enough so that the player does not get bored. This will require the puzzles to have multiple difficulty levels as different people will have different knowledge about certain topics. For the first prototype, the puzzles where not developed to this level as explained in section 5.1. The debriefing and observations in section 5.6 also showed that the players where more focussed on the technology than on the puzzles. The participants did mention they liked the game so it is assumed that the current set of pictures is a good start to develop the puzzles further.

R14: The game needs to incite movement in both the left and the right hand of the player.
To incite movement in both hands, the game concept implements two differently coloured packages (section 4.3). Each type can only be picked up with either the left or the right hand. In the first prototype only the right hand could be tracked and this requirement was not yet tested (section 5.1). Further study will have to show if this solution is the correct one.

R15: The game has to have a gamified scoring system.
To increase the engagement of the player, the scoring system has to be more than a set of numbers. Therefore each correctly sent package will earn the player a collectible post stamp, as discussed in section 4.3. It was chosen to use stamps as it is a renowned type of collectible and because it is close to the chosen setting. The scoring system was not part of the implementation (section 5.1) and later design iterations will have to show if the right design choice was made.

R16: The game needs to prevent demotivating situations from happening.
To prevent demotivating situations from happening, any form of scoring should not emphasise their disability. Also the dynamic package size, discussed in chapter 4.3, will make sure that even if the disorder has progressed since a previous assessment, the movement difficulty will stay the same. From the observations in section 5.6, the participants in the tests were mostly demotivated because of problems with the hand tracking. This shows that this requirement not only influences the concept but is also important for the AR system. It did make it difficult to perceive if the game concept created demotivating situations which should be tested in later design iterations.

R17: The game's tracking algorithms needs to be optimized for both the front and the back of the hand.
When the tracking algorithms are capable of recognising both sides of a hand, the functional movements in the game can be applied normally in the game. This is necessary as otherwise the system cannot conduct the right measurements for an assessment. As one participant in the test could perform a session with both the back and the front of the hand recognised, it should be possible to optimize the system so that it is recognised with everyone.

R18: The game needs to be displayed in full 3D.
The HMD used, described in chapter 5.3, is capable off full 3D. This does require a lot of work as everything has to be generated from two slightly different angles. Developing this is important for the depth perception of the player and should definitely be done in future iterations.

R19: The game has to have an intuitive way to enlarge the image.
As was mentioned in section 5.6, the current solution to enlarge the image was not fully appreciated by the participants. The intuitive action that was observed with some of the participants was to pull the package closer to get a better look. Incorporating this action in the interaction model will create a better solution. This could be done by enlarging the image when the player holds the box within a certain range from the HMD. The problem with this motion is that when the player holds his hand too close, it could obscure the marker from the sensor. Without the marker to which the virtual environment is tethered the environment will not be displayed which could confuse the player. Still, this solution has the most potential as long as the player does not bring the box too close.

R20: The game has to have an encompassing virtual environment with a semicircle sorting box.
Having an encompassing virtual environment is vital when measuring the reachable workspace, as is explained in section 4.1. With the current setup, such an environment is difficult as the sensors on the HMD will lose sight of the marker when looking to the side. A possible solution to this could be to use multiple markers to cover a larger area around the player.

5.8.1. Limitations of the study

The sample group that participated in the experiment was small which diminishes the trustworthiness of this study. The target audience for the designed game are motor dysfunction patients whose only further distinction is that they are usually above 50. The personal characteristics of the participants in the experiment do comply with this. The equal division in gender and the large diversity in type of jobs show a good comparison to the target audience. The one left handed person on eight people is also comparable to the standard human society. The results from the interviews done in chapter 3 show that the current target group is also not that proficient with computers as it only showed up in the latter part of their lives and most of them did not have the interest to learn to use it. The current results show that the participants in the test sessions are a bit younger than the interviewee's and the data supports the idea that younger people have more experience with computers as it has become an unavoidable part of society. It can thus be assumed that the target audience is on average a bit less experienced with computers than the sample group. The same thing applies to their capability for motion. As the sample group are healthy people in contrast to the motor dysfunction that the target group has, they will have less trouble moving and thus controlling the game. This means that any positive result only shows that it is possible that the game works and that it still needs to be tested on patients. It also means that if a part does not work now it will definitely not work with the actual patients. Although all these factors are in favour of a compatible sample group, the sample size is still too low to give definitive answers but it should be enough to get an indication. As the current results already give an unclear answer to the success of this thesis, a larger study has to be performed to verify the approach used and to find the answer to the research question.

6. Conclusion

To answer the research question "How can an AR game that facilitates engaging motion disorder assessment be designed?", several sub questions were asked and answered resulting in a set of requirements that are vital to the game. The first question answered was how games are designed. In section 2.2, two types of approaches where recognised namely the serious game approach, focussing on solving a problem, and the entertaining game approach, focussing on the players. It was chosen that to make a motion disorder assessment game that also facilitates engagement a combined approach would work best. Therefore the approach of Rollings and Adams (2003) was incorporated into the approach of Meijer (2009). The design cycle from Meijer, shown in Figure 3, is a structured set of steps that leads to a finished product while also giving space for the design theories of Rollings and Adams. These theories focus on identifying which design choices have to be made and how these can be made to fit the target audience. The consequences of this approach is that when requirements where elicited in the study, there was a secondary focus on how to improve the engagement of the player. This resulted in player-specific requirements that in turn were implemented in the game so that it can facilitate the engagement of the player.

The second question asked what engagement entailed. Section 2.3 discusses that it can be defined as "Being concerned with all the qualities of an experience that really pull people in and keeps them there" and then continues to use several studies to create a framework for engagement in games. First, there are the seven categories of game experiences that IJsselsteijn (2008) identifies to measure game enjoyment. These experiences represent the feelings a person has when he is engaged in a game. To get these experiences, game design choices have to be made that, following Schell's (2005) research, match the inherent interest of the player, is well presented and has a close psychological proximity. The choices to be made are identified by Rollings and Adams (2003) five game dimensions. The key to this framework is suspension of disbelief. Based on the choices made, the designer creates a world that the player chooses to believe in, which is possible if it follows Schell's three rules. If it goes too far or if things are inconsistent, it can break the suspension of disbelief resulting in the player losing engagement, realizing that it is just a game. This framework was then applied in the design of the game. During Meijer's (2009) #3, common inherent interests and psychological proximities among the target audience were identified. When making design choices, the developed framework was then used to determine how the target audience could be better engaged in the game. Lastly, during the tests in section 5, the engagement of the participants were measured using the game experiences of IJsselsteijn (2008) to determine if the choices made had the required effect.

Another question was what the opportunities and limitations of AR are. AR is a growing technology and current research is focussed on applying it in new fields. Its intuitive control, 3D display, possibility for motion tracking and motion input makes it suited for the motion disorder assessment game. AR also creates the possibility to further enhance assessments with a build-in collaboration function. AR is at the start of commercial use and there are still many aspects being researched. Especially the tracking of real life objects is an area that is still developing but also better display techniques and miniaturization of equipment is still studied. For this thesis, it meant that the game was developed based on the technology now available while taking into account how these may grow in the coming years.

The fourth sub-question consists out of understanding what is necessary to perform the motion disorder assessment. This question is also part of the requirement elicitation in the design approach chosen. The other part of the requirement elicitation is the requirements from the target audience and both are discussed in chapter 3. Based on observations of assessments plus interviews with clinicians and patients, a list of requirements was compiled. To perform the assessment, the patient needs to perform functional movements to the limit of his motion capabilities. The game needs to incite these movements in such a way that quantitative data can be gathered. The player's requires that the game challenges the mind in a realistic setting and does not put any time pressure on the exercises. The requirements gathered are the initial factors vital to the game design. Using the gathered requirements a game concept was then developed to verify and improve this list of requirements.

The fifth question, how can the assessment be made engaging in a game?, is answered by using the knowledge gathered till this point to create a concept game. This concept incorporates the necessary factors for an engaging motion disorder game, which is explained in chapter 4. The designed game is called post office trouble. In the game, the player plays a post office worker that has to sort the international mail which has been encrypted with images corresponding to countries. The player has to pick up a package, identify to which country it has to be sent and then put the package in the correct box. As the interaction is done using the actual motions, the player is performing functional movements that will be measured for the assessment. The image encryption is there to challenge the mind of the player, as this was found to be the best way to engage the target audience. The post office setting was chosen as the target audience preferred a realistic setting which also improves the engagement of the game.

To test if the designed game is truly engaging, which was the last sub-question asked, the game was tested with the help of people in the same age-group as the target audience. Eight participants played an implementation of the game after which they filled in a usability questionnaire and an engagement questionnaire. The results show a large variety in the usability scores and the engagement scores show that although the game did not create negative feelings, the positive feelings scored only average. There are three ways to explain these results.

1. The first possibility is that the game itself is interesting but not engaging. This means that the current factors were not enough to make the game engaging and further research should be performed to improve this.
2. The second possibility is that both the system and the game do not facilitate the engagement of the player but the scores are the result of the players being interested in the novelty of the technology. The engagement scores were thus the result of players being excited to try out AR. This would mean that the factors found have no effect on the engagement and that this thesis shows how not to create an engaging game.
3. The last possibility is that the system does not support the engagement but the game does. The system's low usability is supressing the positive feelings while there are also no negative feelings. This would mean that the factors used where effective and that the game is on the right track.

The main research question "How can an AR game that facilitates engaging motion disorder assessment be designed?" is thus answered with the list of requirements shown in Table 14. These requirements should result in an engaging game that will guide the player in making all the necessary motions for a motion disorder assessment. This thesis could not verify that these requirements truly have the desired effect and further study should thus be performed before they are continued to be used in the TIM project.

This thesis also applied a novel approach to game design in which both serious game theories and entertainment game theories were applied. The approach of Meijer (2009) gave an easy to follow procedure that worked well with the game design theory of Rollings and Adams (2003). Adding the focus on the target audience in the build up to a list of requirements worked well and Rollings' and Adams' game concept worksheet gave a step by step walkthrough where the requirements could be added. This step was a bit chaotic as there was no planning of the requirements. The design process consisted of making a concept, checking the requirements and adapting the concept to fit them. An integration step where the requirements are matched to game elements, like the system component/game element matrix that Duke (1980) uses, is necessary to better integrate the approaches. Without a clear answer to the main research question, it is hard to say if the combined approach was a success. Still, because there were no problems in the approach and it allowed easy incorporation of engagement focussed elements, the approach is considered a possible solution to engagement focused serious game design.

To continue on this research, there are several recommendations:

- First it is recommended to improve the usability of the game. The last set of requirements gives a good direction to the improvements. A professional HMD, instead of two parts glued together, will also help in improving the usability.
- Second, it is recommended to study how the AR environment can be made to encompass the player while still fettering it to a marker or markers. This will allow the sorting cabinet to be placed around the player which will increase the angles of the functional movements made.
- After the first two recommendations, it should be possible to validate the engagement value of the game. A larger test similar to the one performed with the same questionnaires will give better results to determine if the game is engaging or not.
- Lastly it is recommended to study if the game actually improves the assessment method. This can be done in an experiment with an experimental and a control group. The experimental group is assessed using the motion disorder assessment game. The control group performs the normal physical with similar motions. Analysing the results of the assessment will then show if the game adds to the assessment. This experiment can be performed with only the functional game, in which the doctor observes both groups with the qualitative measurements now in use. If the system is capable of data gathering when performing the experiment, the control group performs the normal physical while wearing the HMD for measurements.

Reflection

Looking back at my time in the Systems Engineering group while working on my thesis it has been a difficult period. A combination of personal loss, stress, problems at home and bad planning has made it a long year. My work in developing a first prototype motion disorder assessment game has been one of the brighter elements that also gave me the opportunity to learn a lot. It has shown me how complicated game design can become when dealing with multiple actors, each with their own requirements. Even though lecture after lecture tells you this is a difficult part, you only start appreciating it when you have to deal with it yourself. Being able to apply the lessons learned in these various courses and seeing how they interweave with each other was a valuable experience.

After the challenging and rewarding experience of designing the game, the writing part was very difficult. As usual I started too late with writing. Something that I promise myself not to do after each project but somehow always happens. Wanting to finish the project I forgot all the lessons on scientific writing which resulted in a first version not worthy of the name "thesis". It is a good example of how haste can only make things harder. It thought me to first think about what you are going to do and what you know about it. The knowledge about how to write is just as important as what you are going to write and by forgetting the first it is impossible to properly tell the second. Writing this version of my thesis has been challenging and properly applying scientific writing to an iterative design study has taught me a lot. Still, there is a lot more for me to learn in this area.

All in all, it has been a great experience to develop something that will hopefully help elderly people being treated for their motion disorder. The application of gaming to truly help people is inspiring and I hope to be able to continue doing this in my future.

Bibliography

Abdo, Wilson F., Bart P. C. van de Warrenburg, David J. Burn, Niall P. Quinn, and Bastiaan R. Bloem. 2010. "The Clinical Approach to Movement Disorders." *Nature Reviews Neurology* 6 (1). Nature Publishing Group: 29–37. doi:10.1038/nrneurol.2009.196. http://www.nature.com/doifinder/10.1038/nrneurol.2009.196.

Alamri, Atif, Jongeun Cha, Abdulmotaleb El Saddik, and Abdulmotaleb El Saddik. 2010. "AR-Rehab: An Augmented Reality Framework for Post Stroke-Patient Rehabilitation." *IEEE Transactions on Instrumentation and Measurement* 59 (10): 2554–63. doi:10.1109/TIM.2010.2057750. http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5567156.

Azuma, Ronald. 1997. "A Survey of Augmented Reality." *Presence: Teleoperators and Virtual Environments* 6 (August): 355–85. doi:10.1.1.30.4999. http://scholar.google.com/scholar?q=intitle:A+Survey+of+Augmented+Reality#0.

Azuma, Ronald, Reinhold Behringer, Steven Feiner, Simon Julier, and Blair Macintyre. 2001. "Recent Advances in Augmented Reality." *IEEE Computer Graphics and Applications*, no. December: 34–47.

Bardorfer, Ales, Marko Munih, Anton Zupan, and Alenka Primozic. 2001. "Upper Limb Motion Analysis Using Haptic Interface." *Transactions on Mechatronics* 6 (3): 253–60.

Benyon, D., Turner, P., & Turner, S. 2005. *Designing Interactive Systems*. Harlow, England: Addison-Wesley.

Billinghurst, Mark, Adrian Clark, and Gun Lee. 2015. "A Survey of Augmented Reality Augmented Reality (AR) Defini3on." *Foundations and Trends in Human-Computer Interaction* 8 (2-3): 73–272. doi:10.1561/1100000049.

Botella, C M, M C Juan, R M Baños, M Alcañiz, V Guillén, and B Rey. 2005. "Mixing Realities? An Application of Augmented Reality for the Treatment of Cockroach Phobia." *Cyberpsychology & Behavior* 8 (2): 162–71.

Broeren, Jurgen, Ann Björkdahl, Ragnar Pascher, and Martin Rydmark. 2002. "Virtual Reality and Haptics as an Assessment Device in the Postacute Phase after Stroke." *Cyberpsychology & Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society* 5 (3): 207–11. doi:10.1089/109493102760147196.

Brooke. 1996. "Brooke, J.: SUS: A 'quick and Dirty' Usability Scale." *In: Jordan, P.W., Thomas, B., Weerdmeester, B.A., McClelland, I.L* dustrypp: 189–94.

Brooke, John. 2013. "SUS : A Retrospective." *Journal of Usability Studies* 8 (2): 29–40.

Cidota, Marina A, Rory M S Clifford, Paul Dezentje, Stephan G Lukosch, and Paulina J M Bank. 2015. "Affording Visual Feedback for Natural Hand Interaction in AR to Assess Upper Extremity Motor Dysfunction." In *ISMAR 15' IN PRESS*, 1–4.

Csikszentmihalyi, M. 1992. *Flow: The Psychology of Happiness*. London: Random House.

de Lau, Lonneke M L, and Monique M B Breteler. 2006. "Epidemiology of Parkinson's Disease." *The Lancet. Neurology* 5 (6): 525–35. doi:10.1016/S1474-4422(06)70471-9. http://www.sciencedirect.com/science/article/pii/S1474442206704719.

Dickey, M. D. 2005. "Engaging by Design: How Engagement Strategies in Popular Computer and Video Games Can Inform Instructional Design." *Educational Technology Research and Development* 53 (2): 67–83. doi:10.1007/BF02504866.

Duke, Richard D. 1980. "A Paradigm for Game Design." *Simulation & Games* 11 (3): 364–77.

Feigin, Valery L, Mohammad H Forouzanfar, Rita Krishnamurthi, George A Mensah, Derrick A Bennett, Andrew E Moran, Ralph L Sacco, et al. 2014. "From the Global Burden of Disease Study 2010" 383 (9913): 245–54.

Fireproof Games. 2015. "The Room @ Www.fireproofgames.com." http://www.fireproofgames.com/games/the-room.

Garris, R., R. Ahlers, and J. E. Driskell. 2002. "Games, Motivation, and Learning: A Research and Practice Model." *Simulation & Gaming* 33 (4): 441–67. doi:10.1177/1046878102238607. http://www.scopus.com/inward/record.url?eid=2-s2.0-67549103767&partnerID=tZOtx3y1.

Gibbs, Samuel. 2014. "Google Glass Review@ Www.theguardian.com." *The Guardian*. http://www.theguardian.com/technology/2014/dec/03/google-glass-review-curiously-useful-overpriced-socially-awkward.

Han, Jay J., Gregorij Kurillo, R. Ted Abresch, Alina Nicorici, and Ruzena Bajcsy. 2013. "Validity, Reliability, and Sensitivity of a 3D Vision Sensor-Based Upper Extremity Reachable Workspace Evaluation in Neuromuscular Diseases." *PLOS Currents Muscular Dystrophy*.

Healthcommunities.com. 2015a. "Overview-of-Movement-Disorders @ Www.Healthcommunities.Com." http://www.healthcommunities.com/movement-disorders/overview-of-movement-disorders.shtml.

———. 2015b. "Treatment @ Www.healthcommunities.com." http://www.healthcommunities.com/movement-disorders/treatment.shtml.

Hevner, Alan R. 2007. "A Three Cycle View of Design Science Research A Three Cycle View of Design Science Research." *Scandinavian Journal of Information Systems* 19 (2).

Hevner, Alan R., Salvatore T March, Jinsoo Park, and Sudha Ram. 2004. "Design Science in Information Systems Research." *MIS Quarterly* 28 (520): 75–105. doi:10.2307/249422. http://www.hec.unil.ch/yp/HCI/articles/hevner04.pdf.

Ijsselsteijn, Wijnand, Wouter Van Den Hoogen, Christoph Klimmt, Yvonne De Kort, Craig Lindley, Klaus Mathiak, Karolien Poels, Niklas Ravaja, Marko Turpeinen, and Peter Vorderer. 2008. "Measuring the Experience of Digital Game Enjoyment." In *Proceedings of Measuring Behavior*, 7–8.

Ijsselsteijn, Wijnand, Henk Herman Nap, Yvonne de Kort, and Karolien Poels. 2007. "Digital Game Design for Elderly Users." In *Proceedings of the 2007 Conference on Future Play - Future Play '07*, 17. New York, New York, USA: ACM Press. doi:10.1145/1328202.1328206. http://www.scopus.com/inward/record.url?eid=2-s2.0-70049091260&partnerID=tZOtx3y1.

Intel.com. 2015. "Intel® RealSense™ Camera (F200) @ Software.intel.com." https://software.intel.com/en-us/realsense/f200camera.

Jack, David, Rares Boian, Alma S. Merians, Marilyn Tremaine, Grigore C. Burdea, Sergei V. Adamovich, Michael Recce, and Howard Poizner. 2001. "Virtual Reality-Enhanced Stroke Rehabilitation." *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 9 (3): 308–18. doi:10.1109/7333.948460.

Khademi, Maryam, Hossein Mousavi Hondori, Cristina Videira Lopes, Lucy Dodakian, and Steve C Cramer. 2012. "Haptic Augmented Reality to Monitor Human Arm ' S Stiffness in Rehabilitation." In *International Conference on Biomedical Engineering and Sciences*, 892–95.

Layar. 2015. "Layar @ Www.layar.com." Amsterdam. https://www.layar.com/.

Meijer, Sebastiaan. 2009. "The Organisation of Transactions. Studying Supply Networks Using Gaming Simulation." Wageningen University.

Merriam-webster. 2015. "Engages @ Www.merriam-Webster.com." http://www.merriam-webster.com/dictionary/engages.

Microsoft. 2015. "Hololens @ Www.microsoft.com." https://www.microsoft.com/microsoft-hololens/en-us.

Mozaffarian, Dariush, Emelia J, Benjamin, Alan S. Go, Donna K. Arnett, Michael J. Blaha, Mary Cushman, et al. 2015. "National Health and Nutrition Examination Survey: 2009–2012." *American Heart Association* 131 (29).

Norman, Kent L. 2013. "GEQ (Game Engagement/experience Questionnaire): A Review of Two Papers." *Interacting with Computers* 25 (4): 278–83. doi:10.1093/iwc/iwt009.

Poelman, Ronald, Oytun Akman, Stephan G Lukosch, and Pieter Jonker. 2012. "As If Being There: Mediated Reality for Crime Scene Investigation." In *CSCW'12*, 1267–76. doi:10.1145/2145204.2145394. http://dl.acm.org/citation.cfm?id=2145394.

Rollings, Andrew, and Ernest Adams. 2003. *On Game Design*. 1st editio. New Riders Publishing.

Schell, Jesse. 2005. "Understanding Entertainment." *Computers in Entertainment* 3 (1): 6. doi:10.1145/1057270.1057284.

Shen, Y., S. K. Ong, and a. Y C Nee. 2010. "Augmented Reality for Collaborative Product Design and Development." *Design Studies* 31 (2). Elsevier Ltd: 118–45. doi:10.1016/j.destud.2009.11.001. http://dx.doi.org/10.1016/j.destud.2009.11.001.

Stichting Parkinsonfonds. 2015. "Wat I de Ziekte van-Parkinson @ Www.parkinsonfonds.nl." http://www.parkinsonfonds.nl/over-de-ziekte-van-parkinson/wat-is-de-ziekte-van-parkinson.

US National Library of Medicine. 2015. "Movementdisorders @ Www.nlm.nih.gov." https://www.nlm.nih.gov/medlineplus/movementdisorders.html.

Van Hilten, J.j., B.P.F. Lelieveldt, R.G.H.H. Nelissen, C.G.M. Meskers, J. Marinus, M. Roerdink, L.J.P. van der Maaten, et al. 2014. "Technology in Motion: Innovative Solutions in Motion Research."

Warlow, Charles P. 1998. "Epidemiology of Stroke." *The Lancet* 352: 198–218. doi:10.1016/B978-1-4160-5478-8.10014-4.

Whitton, Nicola. 2010. "Game Engagement Theory and Adult Learning." *Simulation & Gaming* x (xx): 1–14. doi:10.1177/1046878110378587. http://hdl.handle.net/2173/144828.

Wiederhold, B.K., D.P. Jang, R.G. Gevirtz, S.I. Kim, I.Y. Kim, and M.D. Wiederhold. 2002. "The Treatment of Fear of Flying: A Controlled Study of Imaginal and Virtual Reality Graded Exposure Therapy." *IEEE Transactions on Information Technology in Biomedicine* 6 (3): 218–23. doi:10.1109/TITB.2002.802378. http://www.scopus.com/inward/record.url?eid=2-s2.0-0036708414&partnerID=tZOtx3y1.

Appendix A: Patient Interviews

The interviews took place in an assessment room at the neurology section of the LUMC. Seven Parkinson patients were asked by a local doctor if they would like to perform the interview after their assessment and were then guided to the interviewer. Before the interview, a small explanation of the project and its goals was given and the participants were asked if the interview could be recorded. The interview questions were semi-structured so the questions given in the following sections are meant to indicate the topic of the discussion at that point. For the sake of anonymity all patients will be referred to as male and lists will not follow a patient related order. Colours indicate the answer of each individual.

Personal details

What is your age?

- 80 years
- 57 years
- 71 years
- 79 Years
- 74 Years
- 60 years
- 76 Years

What is or was your occupation?

- Primary school teacher
- Journalist
- Car mechanic, later to office in the same company
- Dentist assistant
- Political journalist
- Railroads employee
- Physiotherapist

What is your computer experience?

- Has one and has been interested in them since the beginning
- Only use it for work
- Worked with the PC at work filling out orders but that was 19 years ago. Now he has trouble remembering which makes using the PC difficult. Sometimes plays a game of Patience. Also bought an IPad as an experiment but he often has trouble using it.
- Used at work for billing, never use it at home
- Used it for work, now use it at home for internet and e-mail.
- Only uses it at work or for e-mail
- Use it for work at the end of his career, now uses it for internet and e-mail

Experience with gaming

Did you ever play a computer game?

- Yes, plays angry birds on his phone. Likes the aiming and precision of the game.
- Played a bit in the beginning which influenced his children. He liked how the guy was shuffling through the screen. Never really played much mostly because it did not interest him
- Plays a game of Patience
- Yes, cards and the game with the balls. Likes the challenge and the puzzle elements, although not the ones where you have to put together an image
- Only with the grand-children, but that never works out. Doesn't understand the purpose or function and always has trouble making the system do what he wants it to do.
- No, was never interested
- Bridge, mostly to practice for the real deal.

Did you ever play board or card games?

- Chess, Likes to find out what his level is and wants to keep improving himself
- Sometimes, likes the educative elements in
- Cards, Yatsee, rummicub (with words), Likes the social element[3], being busy, training his memory and challenging his brains.
- Only with the grand-children, never really interested
- Started with chess because of his grandchild, also likes scrabble. Likes the intellectual challenge, keeping his brain fit. Also like the social elements[3].
- Cards, Yatsee, Triomino's , wants to force his brain to work. Likes the challenge and social elements[3].
- Bridge, Rummicub, Card games. Likes how you can plan the many possibilities and choose the best one.

Game element questions

Do you prefer a Realistic or a fictional world?

- Realistic
- Realistic
- Realistic
- Realistic
- Realistic, He thinks this would work better.
- Realistic
- Realistic

[3] The interviewee's used the Dutch word "gezelligheid" which is impossible to translate and therefore the social element is used as it describes what they hint at.

Would you prefer if cultural rules in the game are as realistic as possible, simplified or fictional?

- Simplified
- As close to the truth as possible
- Simplified
- It can be a little bit simplified
- Simplified as adding everything to make it realistic would make the game impossible
- Realistic, Better to understand
- Simplified

Which art style would be preferable (Figure 18**)?**

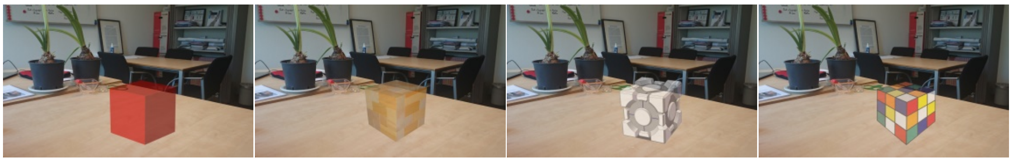


FIGURE 18 - ART STYLE CONCEPTS

- Number 3
- Likes all, especially 3 and 4
- Number 4 because of the colours
- Number 4
- Number 3
- Number 3
- Number 1

Would you like activities to be timed or would you like to take your time?

- Prefers to take his time
- Prefers time pressure
- Prefers to take his time
- Difficult question as he never plays games, likes both options.
- Prefers to take his time
- Prefers to take his time
- Like to get as much time as needed for the exercise

Appendix B – Mock-up

The Mock-up used to present the game concept to the experts exists out of a two-part PowerPoint presentation. The first part consists of four slides explaining the project, the goals, the system setup and an initial game explanation. The second part is a set of images connected through hyperlinks to show how the game looks like and how it works. Below each slide is an explanation of the mode's function.

#	Slide

1

POST OFFICE TROUBLE

A serious game for diagnosing motion disorders

Diagnosing a motion disorder is currently done using a qualitative method meaning that doctors have no hard data for their decisions. To help these doctors the TU Delft and the LUMC are working together on developing a augmented reality system that measures hand movements, thus collecting quantitative data for the diagnosis.

The system includes a game that will incite the patient to make the necessary movements for a diagnosis. I am currently developing the first prototype of this game with the focus on engaging the target audience. This document contains a mock-up of the game and I was hoping you would be willing to give some feedback on it.

2

Post Office Trouble: some info

- Engagement:
  - "To get and keep the interest of someone" (Merriam-webster n.d.)
  - "being concerned with all the qualities of an experience that really pull people in" (Benyon, D., Turner, P., & Turner 2005, p.61)
- Target audience: for the prototype we focus on Parkinson patients
  - Generally older than 50
  - Symptoms include:
    - Tremors
    - Stiff facial expression
    - Difficulty in speech
    - Freezing movements
    - Stiff arms and legs
  - Game length: 10 to 15 minutes (takes place at the hospital)
- Movements needed in game:
  - Grabbing and moving an object
  - The reachable workspace of the patient

3

Post Office Trouble: Setup

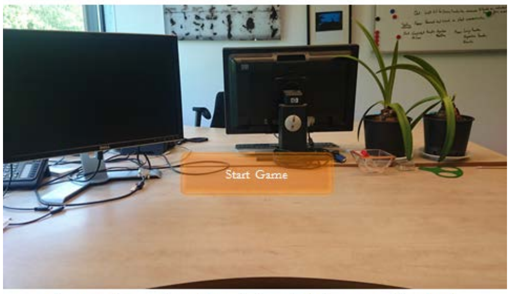
- The patient wears the Cinoptics Airo 2 augmented reality glasses during the game and sits in front of a desk
- Sensors allow for precision tracking of hand movement within visual range
- See-through display overlays images on reality

4

Post Office Trouble: Mock-Up

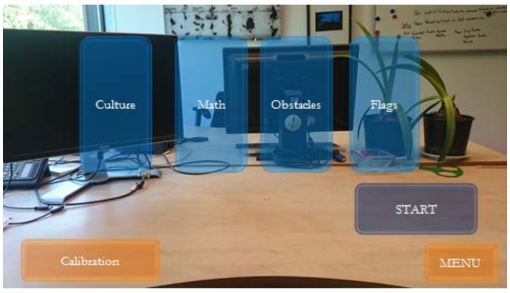
- In the game, you are a post-man after April fool's day. Somebody switched all country names with puzzles and you will have to figure out where to send each package in the sorting station
- The mock-up is fully interactive, so click the button's to "play" the game.

5



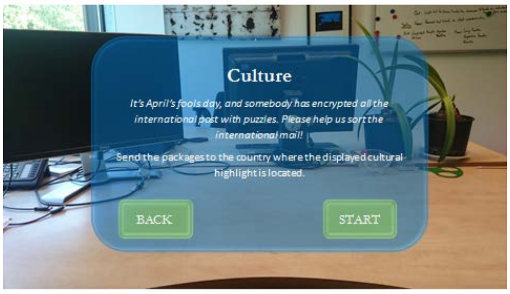
- The image above is the reality, a desk on which the game will be played during the next slides
- The button is virtual, it will be slightly see through since it is displayed on a see trough display.(click the button to start the mock-up)

6



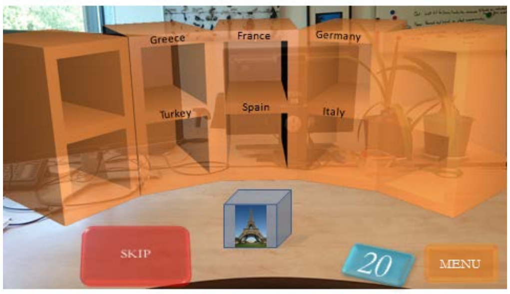
- Main Menu: Use the blue buttons to select the topics you want to play, then press start to play.
- Go to menu -> exit to finish the game

7



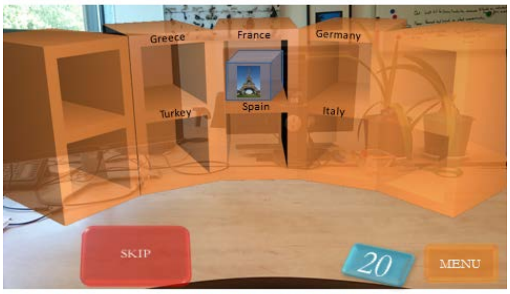
- Short description of the goal of the activated topic

8



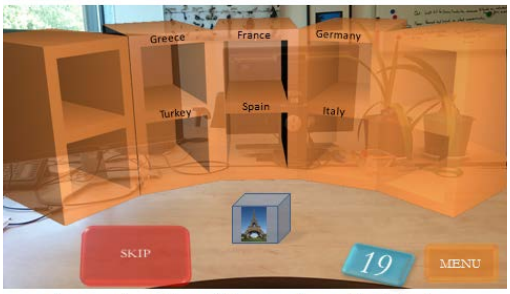
- The game: The blue package can be moved by grabbing it as if it is a real box. The number down on the right indicates the packages left.

9



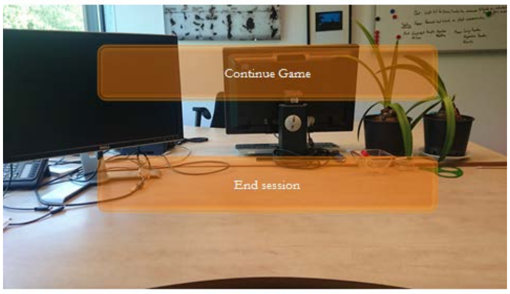
- Placing the package in one of the boxes will finish this package and start the next one

10



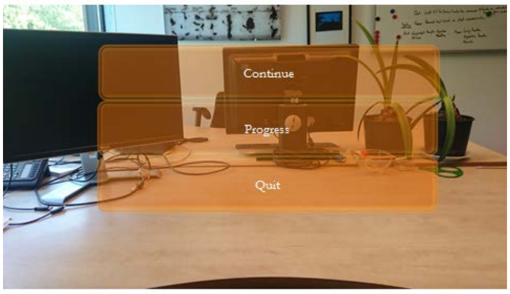
- The next package will have a different size, dynamically calculated (or adjusted by the doctor) for a better diagnosis
- If the patient cannot pick up the package, the skip button can be used

11



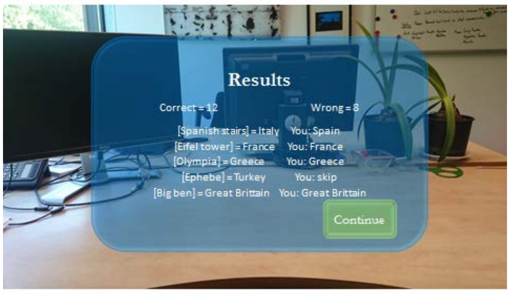
- From here you can end the session to return to the main menu or continue the game.

12



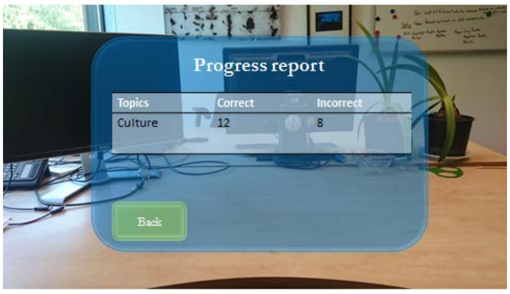
- From here you can quit the game, see your current results or continue playing

13



- After each session, a screen shows the results including the answers of each puzzle.

14



- This displays the current results of each session played.

15

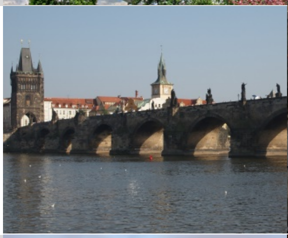
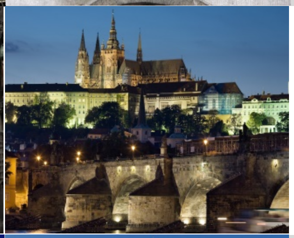
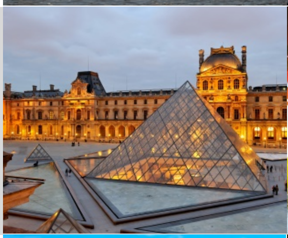
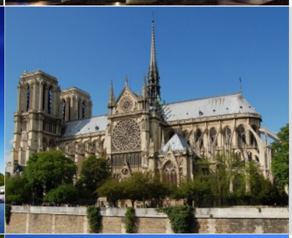
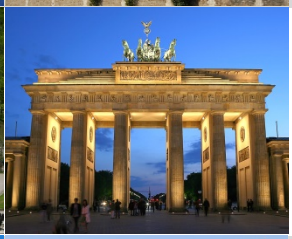
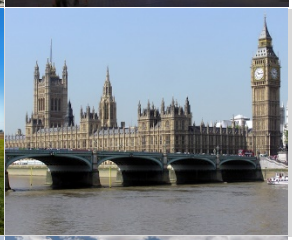
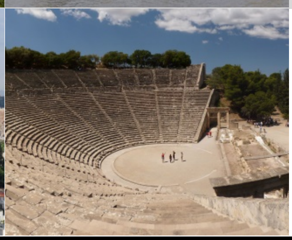
Post Office Trouble: Doctor's side

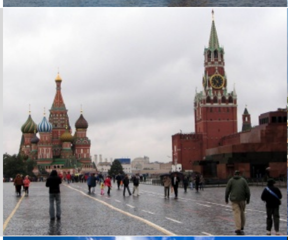
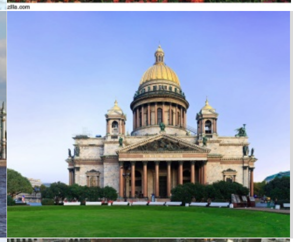
- The doctor influences the game through a PC
- Parameters the doctor can alter are:
  - Amount and size of the sorting boxes
  - Distance between player and boxes
  - Amount of packages per session
  - During session:
    - Size of each upcoming package
    - Puzzle difficulty level
- This to push the patient to the limit of his capabilities so that an accurate diagnosis can be made of his motion disorder

Appendix C – The implemented Puzzles

Two puzzles were implemented in the game. In the puzzle, the player has to recognise the image on the package and match it to the country of origin. In the current version there are ten countries and three pictures per country per puzzle. In the future these numbers need to be expanded to allow for a wide variety in sessions.

Puzzle 1: Matching cultural attractions

Country	Images		
Belgium			
Czech Republic			
France			
Germany			
Great Brittain			
Greece			

Italy	  
Netherlands	  
Russia	  
Spain	  

Puzzle 2: Champion league team Logo's

Country	Image
Belgium	  
Czech Republic	  
France	  

Germany	
Great Brittain	
Greece	
Italy	
Netherlands	
Russia	
Spain	

Appendix D – Questionnaires

This appendix explains how to use the two questionnaires used during the study. For the tests the a Dutch translation was used which is also added.

System Usability Scale (SUS)

English version:

1. I think that I would like to use this system frequently
2. I found the system unnecessarily complex
3. I thought the system was easy to use
4. I think that I would need the support of a technical person to be able to use this system
5. I found the various functions in this system were well integrated
6. I thought there was too much inconsistency in this system
7. I would imagine that most people would learn to use this system very quickly
8. I found the system very cumbersome to use
9. I felt very confident using the system
10. I needed to learn a lot of things before I could get going with this system

Dutch version:

1. Ik denk dat ik dit systeem graag regelmatig wil gebruiken
2. Ik vond het systeem onnodig complex
3. Ik vond het systeem makkelijk te gebruiken
4. Ik denk dat ik ondersteuning nodig heb van een technisch persoon om dit systeem te kunnen gebruiken
5. Ik vond dat de verschillende functies in dit systeem erg goed geintegreerd zijn
6. Ik vond dat er teveel tegenstrijdigheden in het systeem zaten
7. Ik kan me voorstellen dat de meeste mensen zeer snel leren om dit systeem te gebruiken
8. Ik vond het systeem erg omslachtig in gebruik
9. Ik voelde me erg vertrouwd met het systeem
10. Ik moest erg veel leren voodat ik aan de gang kon gaan met dit systeem

Scoring calculation

For each question, the participant fills in how much he agrees with the statement were 1 is strongly disagree and 5 is strongly agree. The total score is then calculated with the formula $SUS\ Total = 2.5 \times \left(\sum(Q_i - 1) + \sum(5 - Q_j)\right)$ wherein Q_i represents the score of the odd questions and Q_j the scores of the even questions. The formula results in a score between 0 and 100 and a positive score is above 67 (Brooke 1996; Brooke 2013).

Game Experience Questionnaire (GEQ)

The GEQ exist out of multiple modules. For this study only the Core module was deemed necessary.

English version:

1. I felt content
2. I felt skilful
3. I was interested in the game's story
4. I thought it was fun
5. I was fully occupied with the game
6. I felt happy
7. It gave me a bad mood
8. I thought about other things
9. I found it tiresome
10. I felt competent
11. I thought it was hard
12. It was aesthetically pleasing
13. I forgot everything around me
14. I felt good
15. I was good at it
16. I felt bored
17. I felt successful
18. I felt imaginative
19. I felt that I could explore things
20. I enjoyed it
21. I was fast at reaching the game's targets
22. I felt annoyed
23. I felt pressured
24. I felt irritable
25. I lost track of time
26. I felt challenged
27. I found it impressive
28. I was deeply concentrated in the game
29. I felt frustrated
30. It felt like a rich experience
31. I lost connection with the outside world
32. I felt time pressure
33. I had to put a lot of effort into it

Dutch version:

1. Ik voelde me tevreden
2. Ik voelde me vaardig
3. Ik was geboeid door het verhaal van het spel
4. Ik kon er om lachen
5. Ik was helemaal geabsorbeerd
6. Ik voelde me vrolijk
7. Ik kreeg er slechte zin van
8. Ik was met andere zaken bezig

9. Ik vond het saai
10. Ik voelde me zeker
11. Ik vond het moeilijk
12. Ik vond het aansprekend van vormgeving
13. Ik vergat alles om me heen
14. Ik voelde me lekker
15. Ik was er goed in
16. Ik voelde me verveeld
17. Ik voelde me succesvol
18. Ik voelde me fantasierijk
19. Ik had het gevoel dat ik de gamewereld kon exploreren
20. Ik genoot ervan
21. Ik was snel in het bereiken van de doelen in de game
22. Ik was geïrriteerd
23. Ik voelde me onder druk gezet
24. Ik was prikkelbaar
25. Ik was mijn gevoel voor tijd kwijt
26. Ik voelde me uitgedaagd
27. Ik vond het indrukwekkend
28. Ik was ten volle geconcentreerd op de game
29. Ik was gefrustreerd
30. Ik vond het een rijke ervaring
31. Ik was weg uit de buitenwereld
32. Ik voelde tijdsdruk
33. Ik moest er veel moeite in steken

Scoring calculation

The participant gives scores the statements above with “not at all”(1) or “extremely” (5). The GEQ divides the scores in seven categories. For each category, the average of the scores of several questions is taken. The questions per category are:

- **Competence**: Questions 2, 10, 15, 17, and 21.
- **Sensory and Imaginative Immersion**: Questions 3, 12, 18, 19, 27, and 30.
- **Flow**: Questions 5, 13, 25, 28, and 31.
- **Tension/Annoyance**: Questions 22, 24, and 29.
- **Challenge**: Questions 11, 23, 26, 32, and 33.
- **Negative affect**: Questions 7, 8, 9, and 16.
- **Positive affect**: Questions 1, 4, 6, 14, and 20

Raw Data

Table D1 to D4 show the raw data results of the questionnaires and the calculated final scores:

TABLE D 1 – SOCIAL DATA

ID	Age	Gender	Primary hand	Occupation	Computer experience	AR Experience	Affinity for technology	Digital game experience	Type of platform	Analog game experience
1	61	female	Right	ondersteuning o&s	normal	never	neutral	every day	mobile device	never
2	52	female	Right	Lerares	a little	never	neutral	every day	mobile device	Sometimes
3	79	male	Right	Civiel engineur	a little	never	neutral	never	none	Sometimes
4	54	female	Right	Schade adviseur	normal	never	agree	weekly	mobile device	Sometimes
5	67	male	both	Fysio-therapeut	normal	never	neutral	never	none	once a month
6	66	male	Right	brandweerman	a little	never	disagree	never	none	Sometimes
7	52	female	Left	Universitair docent	much	never	neutral	never	none	Sometimes
8	53	male	Right	Werktuig-kundige	a lot	never	agree	weekly	computer	Sometimes

TABLE D 2 - SUS ANSWERS AND RESULTS

ID	SUS1	SUS2	SUS3	SUS4	SUS5	SUS6	SUS7	SUS8	SUS9	SUS10	SUS_total
1	1	4	1	4	2	3	4	3	1	4	27.5
2	1	2	2	3	3	1	2	3	2	3	45
3	2	1	3	5	3	1	5	1	2	3	60
4	3	1	4	1	4	1	5	1	3	2	82.5
5	3	3	3	3	3	3	3	3	3	3	50
6	3	1	3	3	3	1	2	2	2	5	52.5
7	2	2	4	2		1	5	2	4	1	75
8	3	1	4	1	4	1	5	1	5	1	90

Table D 3 - GEQ answers

ID	GEQ1	GEQ2	GEQ3	GEQ4	GEQ5	GEQ6	GEQ7	GEQ8	GEQ9	GEQ10	GEQ11	GEQ12	GEQ13	GEQ14	GEQ15	GEQ16	GEQ17
1	3	1	2	4	2	2	1	1	1	2	4	4	2	1	1	1	1
2	3	3	2	3	3	3	1	1	1	2	2	2	4	3	2	1	3
3	3	3	5	2	5	3	1	1	1	3	3	5	5	4	2	1	3
4	2	2	4	4	4	1	1	1	1	1	4	4	1	3	1	1	1
5	3	3	1	2	2	3	1	1	1	4	3	1	3	4	3	1	1
6	2	1	4	3	3	2	1	1	1	2	4	4	4	2	1	1	1
7	4	3	4	1	3	3	1	1	1	4	1	4	4	4	4	1	4
8	4	3	4	5	3	5	1	1	1	4	2	4	3	5	3	1	3

ID	GEQ18	GEQ19	GEQ20	GEQ21	GEQ22	GEQ23	GEQ24	GEQ25	GEQ26	GEQ27	GEQ28	GEQ29	GEQ30	GEQ31	GEQ32	GEQ33
1	1	1	3	1	1	1	1	1	4	4	5	4	3	1	1	5
2	1	1	3	2	1	1	1	4	3	2	5	2	3	3	1	2
3	2	1	3	2	1	1	1	5	3	4	5	2	4	5	1	4
4	1	1	3	1	3	1	1	3	4	4	5	3	4	3	1	3
5	1	1	3	1	1	1	1	1	3	3	4	1	3	3	1	3
6	1	2	3	1	1	1	1	2	3	4	4	4	4	4	2	4
7	1	1	3	4	1	1	1	2	4	1	4	1	2	3	1	2
8	3	1	4	4	1	1	1	2	4	5	4	1	4	3	1	2

TABLE D 4 - GEQ RESULTS

ID	Competence	Immersion	Flow	Challenge	Positive affect	Negative affect	Tension
1	1.2	2.5	2.2	3	2.6	1	2
2	2.4	1.8	3.8	1.8	3	1	1.3
3	2.6	3.5	5	2.4	3	1	1.3
4	1.2	3	3.2	2.6	2.6	1	2.3
5	2.4	1.7	2.6	2.2	3	1	1
6	1.2	3.2	3.4	2.8	2.4	1	2
7	3.8	2.2	3.2	1.8	3	1	1
8	3.4	3.5	3	2	4.6	1	1