Implementing Virtual Reality in the Council of Coaches system

Rens van der Werff

Bachelor Thesis Creative Technology

Supervisor: Dr. R. Klaassen

Critical Observer: D.P. Davison

UNIVERSITY OF TWENTE

Abstract

As the population ages the demand for personalised healthcare coaches grows larger. To keep up with this demand, the Council of Coaches (COUCH) was created, aiming to provide personalised coaches from at home, by means of an interactive conversation with some coaches. In order to keep the users using the system and its advice credible, user engagement is important. One way of possibly increasing user engagement is by implementing Virtual Reality (VR) into the system. To find out what features work well in VR in COUCH, two prototypes were created that differentiated in seven different areas, from location to interaction. These two prototypes were each tested by using 6 participants that served as proxy users and 2 target group users that were approached online with an interactive video in order to adhere to the COVID-19 guidelines. The results of these tests showed that the two features who have the most impact on user engagement are the environment, a cosy room worked best in this research, and accessibility, here shaped as subtitles to support the spoken text. A small sample size means that more research on the topic is recommended and more research with the target group should be performed.

Table of contents

Chapter 1: Introduction

1.1 Problem

The world population is slowly ageing as fertility rates continue to decline and life expectancy increases [1]. With more of the population requiring attention from the healthcare sector, a problem is quickly rising. Many older adults are living under the effects of chronic conditions and are in need of coaching in order to cope with this [2]. Unfortunately, there are not enough workers in healthcare to keep this personalized coaching service running. So far, there is still no easy way to provide this without a human being; most systems and tools either focus on a single domain or younger target group, or fail to keep the users engaged [2].

The internet holds another solution to the problem, but has some of the same shortcomings. It has all the information and answers one might need about their health and wellbeing related questions and concerns, but it can be difficult to find reliable information in a bloated web. For older adults this could be even harder, as they are, generally speaking, less fluent with computers and the internet than younger generations. To have one trustworthy place where all this information is easily and always accessible could prove very useful, especially for older adults. That is where the Council of Coaches comes in.

1.2 Council of Coaches

The Council of Coaches [3] (from here on named COUCH) is a radically new virtual coaching system consisting of multiple embodied conversational agents (ECA) that can advise the user about all kinds of health and wellbeing concerns. It aims to push the state of the art conversational agents and the interaction with the user and between other agents. The agents are embodied, which means they have a visual representation (humans in COUCH's case). They are used as they provide significant value over normal text replies [4]. The system consists of multiple agents, each with their own expertise, personality and style of coaching, that can both each interact with the user, as well as with each other. Their expertises can include being a social coach, an activity coach, a dietary coach and many more. The user can ask a question or explain a problem to which the coaches will try to find a solution or suggestion by discussing their ideas and opinions with each other and the user. The coaches can suggest a range of solutions, from diets to exercises and are always available, making COUCH much more convenient and cheaper than a human coach.

Figure 1: The Council of Coaches

Having multiple conversational agents who also communicate with each other is what makes COUCH stand out from other solutions. It has also been found that multiple agents increase the credibility and persuasion of a system [5]. This is important for COUCH, as the system tackles health related issues and thus needs to feel believable and reliable. The conversations should also feel genuine and the solutions credible. The target group of COUCH is older adults, as they can benefit the most from such a system. Therefore this paper will also focus on older adults as the target group.

1.3 Motivation

One way that COUCH has not been fully explored yet is the medium through which the information is conveyed to the user. COUCH currently uses a screen and speakers/headphones when available. Although a screen does do everything COUCH needs it to, it is interesting to study the effects alternatives might have. One of the technologies that could be used instead of a normal screen is Virtual Reality (VR).

Virtual Reality or VR is a technology that immerses the users by putting them in a 3D environment which they can explore. The position and rotation of the head and hands is fully tracked, making the person feel as if they were actually there. VR is currently used in many fields, most notably in gaming and educational systems. It can be a great tool to train people what to do in certain situations, without them having to go there in person.

VR thus offers an angle of immersion and engagement a normal screen cannot [6]. It is also better at conveying educational information and the information is remembered longer by its users [7]. Engagement is important for a system that communicates information and it is therefore interesting to see whether VR can improve on this and/or add other benefits, such as the perceptiveness of credibility.

When looking at the target group, VR is not something many older adults have had any experiences with. This does however not mean older adults and VR do not match, on the contrary, VR seems to show added benefits across the board compared to a screen [8]. As VR content can often be fast and flashy, COUCH should make sure it is also comfortable for older adults to use.

1.4 Goal

The goal of this paper is to transform COUCH from being screen based to VR based. COUCH is currently designed and optimised for the use of a monitor and mouse, which means the system will have to change as a whole. This might include a full rework of the interaction system and the possible addition of new interaction mechanics. The environment which COUCH is currently set in will also be changed to one that complements the use of VR better. This is all done with the target group in mind and should thus be tailored to them. Once finished, the VR solution will be tested against the screen based version in order to answer the research question mentioned below.

1.5 Research questions

The main research question that can be formulated from this is:
"What is the effect on user engagement when implementing VR in the Council of Coaches system?"

To answer this question, three sub-research questions have been formulated that will be answered using literature in chapter 2. These questions are:

1. *"What are the benefits/downsides of using VR in applications with ECAs?"*
2. *"What are the VR applications/features that work best for older adults?"*
3. *"What are the best ways to interact with a VR system using conversational agents for older adults?"*

Chapter 2: State of the art

2.1 Introduction

VR has been rising in popularity ever since it was first introduced. Both the entertainment industry and the ‘serious’ industries have adopted this technology. The most popular and most known version of VR is the ‘head mounted display’ (HMD). This is a headset that a user can attach to their heads, which has two screens inside, one for each eye. It is able to track the user’s head and body movements either by using infrared base stations that calculate the position (such as the HTC Vive), or by using cameras on the headset itself, which can track the world around you (such as the Oculus Rift S). Many HMDs are not standalone and need a VR capable computer to run the programs for them. Some are standalone (such as the Oculus Quest), which allows for greater mobility as there are no wires attached, but they need to be charged and are far less powerful, making them unusable for heavy applications.

In the entertainment world the gaming industry is currently seeing much more VR adaptation as the new HMDs are getting very affordable and AAA game developers are now starting to produce games specifically made for VR [32]. The technological advancements made in the gaming industry are then also used in serious applications, such as anxiety therapy [33] and visa versa. The big technical hurdles of the past are gone and developers can make the application they want.

2.2 Related Work

COUCH is developing a product that tries to alleviate some of the problems in the current world. There are, however, more companies and institutions that are doing or have done, the same or similar things. In order to set COUCH apart from its ‘competition’ and to see what is currently possible as proven by other products, it is important to get an overview of all the related work and current solutions that have some similarities to COUCH. With this information, a conclusion can be drawn about whether COUCH is novel and its goals are feasible.

The ECAs of COUCH can be compared to many intelligent conversational agents out there. Smart assistants such as Google Assistant, Siri, Alexa and many more all related (somewhat) to the agents of COUCH. But while these smart assistants try to be good at everything, the agents of COUCH are more specialized and embodied. Embodying the agents is done to humanize the agents and also because giving the agents an avatar or visual representation of a human can increase user engagement and enrich virtual spaces [9]. Giving agents non-verbal behavior has also shown to increase engagement and conversational contributions [10].

2.2.1 Nina - Nuance

Nina [19] is an intelligent virtual assistant designed to work on webpages, apps and many more and can provide an automated experience for customers by engaging them in natural conversations. It is a smart assistant that helps the customers make choices and has won many prizes for its competence. Unlike COUCH, Nina does not have a visual representation and looks more like a traditional chatbot. Nina can work for many companies and can adjust quickly to their image, tone and products and can learn from her interactions with customers. This means she is constantly evolving and optimizing her behavior to improve the accuracy of responses.

Despite this system not being too similar to COUCH, Nina is also pushing the state of the art of conversational agents. She provides intelligent, human-like and refined conversations that can engage and persuade the customers at the right time. She also understands complex questions very well, something COUCH can not yet do. In order to improve the conversations of COUCH, the agents should be improved so that they, like Nina, allow complex answers and questions and can respond correctly.

Nuance also has other solutions currently working in the healthcare industry, but does not (yet) combine the healthcare solutions with their intelligent agents as COUCH is trying to.

2.2.2 Ada and Grace - Museum of Science, Boston

Ada and Grace [20] are an example of intelligent virtual agents that converse with museum visitors. They are also visualized using computer generated character animation and displayed to look life-sized and life-like. The agents speak to the visitors directly and can tell, answer all kinds of questions about the contents of the museum exhibits, suggest exhibits to visit next, but they can also be funny and display a range of human emotions.

Figure 2: Ada and Grace; source: Museum of Science Boston

Because the agents are using cutting edge technologies and are an intriguing display of advanced STEM (Science, Technology, Engineering and Mathematics) technology, they themselves are a technological exhibition as well. The ‘Science Behind Virtual Humans’ exhibit is a dynamic exhibit that educated the visitors by showing the underlying processing of the agents, such as automated speech recognition and natural language processing, that make the agents feel like real people. This is all done in an effort to inspire youth and learners about all kinds of STEM fields.

This system seems very similar to its approach as that of COUCH, but there is a difference. Apart from the area of science they operate in, the two agents seem to interact with each other while in actuality they just add on one another. The answers are scripted and thus so is the interaction between the agents and the user. This means the system acts more like a single agent, but is given two faces in order to humanize the agents more and keep the conversation interesting. While this is different from what COUCH is doing, it is a good example of how conversational agents can be used in an educational way and keep the users’ attention when explaining and advising. It also shows that having more than one agent (even if only seemingly) can work and make conversations more interesting for users.

2.2.3 vHeath - Aetna

With an ever increasing demand for primary healthcare around the world, technological advancements in the field must be made. The need for GPs (general practitioners) in developing countries is rising, where they often do not exist or are out of reach due to financial or geographical reasons. With a GP model being considered an incredibly valuable service, it needs to be available to all people. This is where virtual healthcare comes in. Ofcourse, it cannot replace the current system, but will certainly become an essential part of it. There are many benefits of such a system being introduced, such as a potential reduction in costs, farther reaching healthcare and more efficient patient managing.

Aetna thus developed their own virtual healthcare service vHealth [21]. This application would work best in developing countries where GPs are scarce and would provide many of the services a GP might. A client could, for example, have a video consultation with a virtual Aetna doctor through their mobile phone. The virtual doctor could then make an assessment of the clients symptoms/problems and recommend what to do next. This could be a recommendation of a diet or some prescription, but the system is also able to arrange a specialist that will visit the patient at home for further tests. These test results could then appear online instantly, allowing the virtual doctor to make an appointment at the hospital is necessary. The application can be used on many devices, most notably a smartphone, meaning the application can be used anywhere (if supported by the Aetna network) and a person would never have to leave their home to receive primary healthcare.

While many of the above stated features are currently available in their app, the intelligent agent part is still in development. However, as the core features, such as prescriptions and appointments are already in place, it is easy to see what the system will look like with the agent implementation. After this the system should become more user friendly, especially for those who have difficulty getting around mobile applications, as they can now talk to a ‘person’ that does everything for them.

While the system uses a different technique than COUCH, there are some similarities. Both applications provide some kind of advice in the health sector, with vHealth trying to replace or function like a GP and take over as much medical care as they can, while COUCH tries to offer more lifestyle type advice. As these applications are dealing with health related problems and sensitive user data, they face similar ethical dilemmas, like what an agent can say and recommend.

2.2.4 Conclusion

In conclusion, COUCH appears to be a niche product that targets an area of healthcare that is currently facing problems. Not many comparable commercial systems could be found that try to solve the issue of insufficient coaching for many older adults, as COUCH is trying to. The biggest shakeup to the current solutions is the fact that COUCH uses multiple agents that can also interact with each other. This has not been well implemented yet, at least in a coaching solution. VR adaptation currently also seems to lack in commercial coaching systems, though plenty of research is available. It therefore shows COUCH is a novelty in its area and truly does push the state of the art.

2.3 Relevant Literature

In order to answer the sub research questions as well as possible, the current literature surrounding COUCH's aspects should be reviewed. While this gives insight in the literature, sometimes interviews and observations are needed to get the complete picture. The literature should, however, give a good overview of the advantages and disadvantages of the technologies used by COUCH.

2.3.1 Virtual Reality

Virtual Reality is seeing a boom in popularity and is getting adopted in many areas of innovation, from surgical solutions in hospital [11], to entertainment, to military exercises [12]. This is because there are many reasons VR can provide added benefit over other technologies. Lele [12] names five benefits that are applicable to military purposes, but they also apply (to some extent) to other areas of innovation that use VR, in this case COUCH.

Figure 3: Surgical procedure in VR; source: digitalheatlh.net

The first benefit is that VR makes it possible to simulate near real scenarios, easily adaptable and is able to produce different and specialized tools quickly. With the technology constantly improving, this will only be done quicker and with higher quality. New technologies, such as photogrammetry, allow the use of photos to automatically generate a 3D reconstruction of an object or area, as showcased by Vajak and Livada [13]. This makes it possible to very quickly transform an area in the real into a virtual one. Because VR feels near real, this can be a great alternative to many real life educational scenarios in dangerous professions, such as the mining industry where Pedram et al. [22] found it was well received.

The second benefit is the fact that VR can be a good way to offer answers into how technologies should evolve. Lele [12] states that modern military challenges are both conventional and asymmetric in nature, which demands the need for innovation of current technologies used by the military. VR can be a way to analyze and test new innovations, before the actual production of them. This can provide a clear insight into the necessity and feasibility of new innovations.

The next benefit is that VR is a cost-effective alternative to using real scenarios. This is most notable in areas such as the military and healthcare, where equipment is very expensive. Much of this equipment is one time use only, think of bullets, fuel, syringes or mouth caps and some equipment could get damaged, like expensive medical machines. VR does not cause any loss or damage to equipment, making it far more cost-effective. In many areas of exercise, the possibility also exists of human loss or injuries, such as extreme sports (parachuting, surfing) or high risk professions like flying a plane. In this case VR offers a safe and controlled environment where there is more room for mistakes.

The fourth benefit is the price of the VR headsets and computers themselves. Commercial headsets like the HTC vive or Oculus Rift have become increasingly affordable and VR-ready computers have gone from supercomputers to compact laptops. The accessories of these headsets have likewise gone down in price, making VR as a whole more accessible, not only for companies, but also for consumers.

The final benefit is that VR provides a clear environment that only makes the necessary information available when needed. Lele [12] claims the removal of this 'clutter' allows the decision maker to make more correct, timely and quick decisions during operations.

Although many benefits are already mentioned, the most important ones to this thesis are the immersion of VR and the engagement that it brings with it. As the person is fully inside a virtual world, without any outside distractions, focussing on a given task, the engagement is much improved compared to using a 'normal' screen/monitor [6,7,8]. Allcoat and von Mühlenen [7] add to that by saying that using VR for educational purposes can improve understanding and learning. Improved engagement and learning experience could be due to VR promoting active learning (interactive), whereas watching a screen is considered passive learning, and an increased level of immersion. Allcoat and von Mühlenen [7] also found that VR had a positive impact on the user's mood, increasing positive emotions and decreasing negative ones.

An improved level of immersion means that VR can be used for many 'serious' applications and it has been proven to be beneficial to many applications such as stress therapy [14], motor rehabilitation [15], appendectomy surgery [16] and mental health problems [17]. Holden [15] states that the potential of

VR in healthcare is great, but the cost is too high. However, this paper was published in 2005 when VR was still a very new and expensive technology, but this is no longer the case as noted by Lele [12].

It can be concluded that VR has an abundance of benefits attached to it, however, this does not mean there are no downsides. Although most papers found seem to only provide benefits, Varela-Aldás et al [18] found that exercises in VR generally provide a lower heart rate than normal exercises. This can be linked to physical activity, concluding that exercises in VR are less effective than normal ones. This is not a problem for COUCH however, as the use of the application does require the use of physical exercises.

Probably the biggest downside of VR is that it can cause motion sickness for some people, or in certain scenarios. The interactions the user had with COUCH are simple (in terms of VR), because the user sits down and only has to move their head (and possibly arms). This means the likelihood that the system induces motion sickness is very low compared to systems where the user has to use their whole body and walk around. Another factor that can affect the user's state is the duration of interaction with a VR system, where longer sessions correlate with an increased chance of motion sickness. The interactions with COUCH are considered short interactions, thus this is unlikely to be a problem.

2.3.2 Older adults

Even though VR has plenty of benefits attached to it, these benefits do not necessarily affect all ages to the same effect. Most of the tests conducted in the aforementioned articles are conducted with younger people and not with older adults, the target group of COUCH. The reason for this could be that older adults are not the target group for most applications, but also that older adults generally have more difficulty completing tasks in VR than younger people, such as wayfinding [23], and thus make for harder test subjects. This might be the reason why literature does not show conclusive evidence VR solutions work well (or better) for older adults as it does for younger people. The literature available shows inconsistent findings: some studies show that VR provides little to no benefits over conventional methods [24, 25], while others show it does yield a significant improvement [26].

This shows that whether VR provides benefits to a system is mostly dependent on the system itself. To conclude whether VR would work in COUCH, actual user tests with the target group would seem to be the only option for conclusive evidence. But while the literature found shows mixed results, it does not claim VR provides worse results than conventional methods. This could mean that proper implementation of VR in COUCH should provide equal or better results than using the screen based version, which would still make it a viable alternative.

2.3.3 eHealth

With the increasing popularity of VR, and the global move towards eHealth [27], many applications have risen that try to combine the two. VR has become a good substitution and addition to the current healthcare system as it is getting more recognition and approval from healthcare professionals, such as therapists [28]. Many kinds of cognitive therapy show that VR provides similar results to 'normal' approaches [29], making it a viable alternative. This is confirmed by Turner and Casey [30] who found that VR shows considerable promise in psychological interventions and makes for a good substitute for face-to-face therapy. Lindner et al. [28] also state that costs are no longer a barrier of entry, as VR has been undergoing rapid development and has become significantly more affordable the last few years, as stated before by Lele [12], making it easier to adopt in healthcare practices.

This shows promising signs for COUCH as it aims to take some of the workload of health and wellbeing coaches. This means it needs to provide a similar experience to what a specialist might, which should be possible, considering the aforementioned literature.

2.3.4 Summary

From the literature found a few things can be concluded: VR has many benefits and some downsides, but they can be prevented/reduced with proper implementation and careful design; older adults experience the benefits of VR to a lesser effect, but the right application/implementation can still work well for them; VR shows great potential in eHealth applications. Implementing VR thus seems to be a straightforward advancement for most systems, but it is not guaranteed to be viable. It needs to provide enough benefit in order to warrant the extra costs associated with it. This means that the viability of VR is dependent on the application and the implementation of VR as its component, which at COUCH comes down to whether the immersion and interaction is improved enough by using VR over using a screen.

2.4 Conclusion

When looking at the literature and the related work, enough information is gathered to (at least partly) answer the sub-research questions formulated in section 1.5 in the introduction.

"What are the benefits/downsides of using VR in applications with ECAs?"

VR provides many benefits over using a screen in all different kinds of applications. Most notably, VR provides an increased immersion, which influences a user's engagement and ability to learn and observe. This can prove to be especially useful for a system that uses ECAs to provide important information to the user. ECAs would appear more lifelike and keep user's attention better in VR, while VR improves engagement, which means the users would listen to the ECAs better and thus remember the given advice of the coaches better. The only real downside would be that the system could cause motion sickness, especially for older adults, but this is unlikely to be the case as the movement in COUCH would be very limited and the sessions would not take too long.

"What are the VR applications/features that work best for older adults?"

As mentioned above, having limited movement and short sessions would most likely work best. This should be combined with a simplistic and easy to understand interaction system, so that the barrier of entry older adults can experience with new technologies, is lowered. The interaction should feel intuitive and logical, while the controls in VR are often not a problem due to them being so intuitive.

"What are the best ways to interact with a VR system using conversational agents for older adults?"

While not much literature or works were found on this subject, it could be assumed that older adults would either prefer to speak to the system, or choose out of select options. Speaking to the system would be most intuitive for them and would not require any prior technological knowledge. Sadly, voice recognition is (not yet) possible in COUCH, but it is possible to select an option. This could be due to the fact that COUCH has been developed with the target group in mind.

Chapter 3: Methodology

With the background research of the project completed, a ‘problem’ is stated and the next step is to work towards a solution. This is done in a structured and iterative way by using the Creative Technology Design Process [31], with some small changes to fit the project better. This process is explained below, as well as an outline for the rest of the report.

The Creative technology Design Process is a design process specifically designed for the Creative Technology bachelor at the University of Twente. It is based on two existing models, the Divergence and Convergence models and Spiral models and is visualized in figure 100 below. It aims to add structure to a (sometimes) complex design process and stimulating iterative and flexible ideation/prototyping. The process consists of four phases; Ideation, Specification, Realisation and Evaluation, each phase in the form of a chapter. These phases are described below as they are applied in this thesis.

<u>Ideation</u>

The first phase is all about exploring possible solutions for the problem stated in chapter 1 and 2. These ideas can be found by brainstorming or by looking at similar technologies and adapting them into this project. These ideas will be filtered based on research from literature and related work and a list of ideas to continue with will be chosen. This differs somewhat from the chosen model as the filtering of ideas mostly happens there in the specification phase with the use of quick prototyping. However, due to the nature of VR applications, quick prototyping is more difficult which makes existing research more reliable.

<u>Specification</u>

In the second phase the ideas from the first phase are further elaborated and their technical side will be explored. The user and system requirements are established, such as what kind of technologies are needed to translate the idea from paper to VR (in this case), which provides a clear goal for the realisation phase.

<u>Realisation</u>

This phase describes the process of building the actual prototype from previously established ideas and requirements. Certain aspects will be described more extensively to provide context and difficulties and/or shortcomings will also be discussed.

Evaluation

In this final phase the created prototypes are tested with users in order to answer the main research question. The prototype is reviewed to see whether it has achieved all of its requirements. After the user tests the results will be gathered and organised, so that they can be analysed in the conclusion.

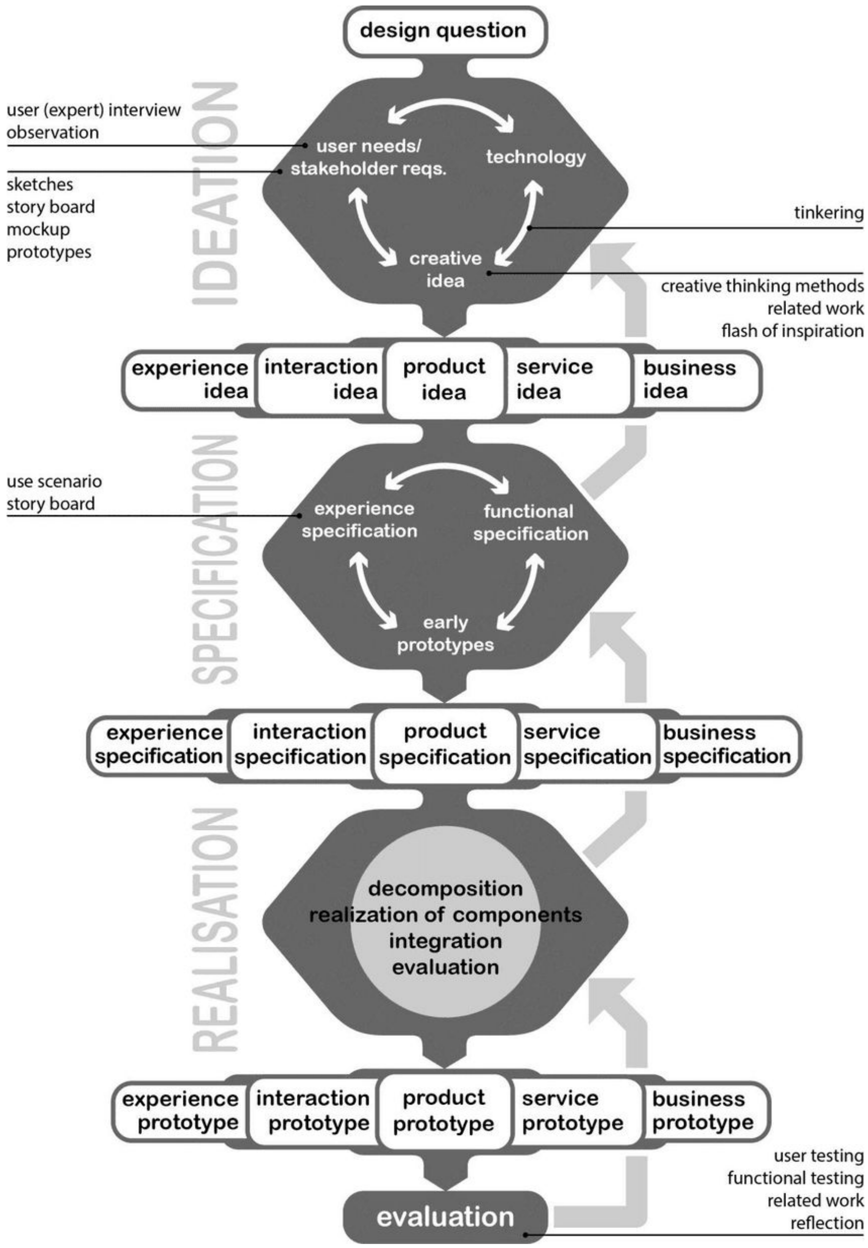


Figure 4: The Creative Technology Design Process

The final chapters of this thesis will be the conclusion and discussion, as more commonly found in research. These chapters will conclude the gathered results and discuss the research process, respectively.

Chapter 4: Ideation

This chapter will take a look at the different possibilities and features when implementing the system proposed in chapter 1. The ideas discussed in this chapter were created in brainstorms by the researcher or added later on when other options became apparent. These were short brainstorming sessions in which the researcher wrote down every idea they could think of. A full overview of the ideas can be found in appendix A1. To add more structure and clarity to the generated ideas, they were all assigned to one of three components of the system. These are components of the system that will be affected by this project and thus each have multiple solutions. These solutions will be compared to each other, as well as with the current implementation, to see which would be best suited for the application and the target group. The three components of the system that will be ideated are:

- (Graphical) User Interface
  The interactive elements in the system.
- Interaction
  The way the user interacts with the system
- Visualisation
  The appearance of environmental elements

Although these components try to divide the system to make it easier to ideate and find existing material, the sections are highly intertwined, which will cause some overlap in ideation between them.

Interviews or paper prototypes are often used in research to get an impression of the target groups' response towards certain ideas. While this can be a great tool for many kinds of research, it is less fitting for this thesis. In order to get usable information from these tests, the participants (older adults) would be asked to compare different concepts of system solutions and provide their opinion and preferences. However, without an actual working prototype, the concept solutions would appear too abstract for the participants, as most of them have little experience with computer interaction and almost no experience with VR. Therefore, the choices of which features will be realised will be based on literature, existing systems, and guidelines set up by major innovators in the VR area.

4.1 User Interface

While the user interface (UI) is greatly dependent on the way the user interacts with the system, some rules/guidelines can be established that affect every version of the system. The system should of course focus on being user friendly and the user experience (UX) is paramount. There are certain guidelines when it comes to making a user interface that can assist in making a successful one, as described by the U.S. Department of Health & Human Services [34] for example. With these guidelines and older adults as the target group in mind, the following points can be formed and should be followed when realising the system:

- UI should be responsive and intuitive
- Text should be clearly readable (font and size are important)
- Buttons should be large and easily accessible
- Different options should be easily distinguishable (content, colour and position are important)
- The user should be provided sufficient time to read and consider different options

Many of the UI features are already set in place by COUCH, making the design process of the VR integration more simple. While in conversation with the various coaches, the user is asked several questions and prompted to select the most fitting answer out of three options. This is because the current version of the platform does not support language interpretation and thus relies on pre-scripted options. The questions that are asked are formulated to fit this feature. This means the system does not allow for custom answers written on a keyboard or by voice recognition. The newest version of the system now also supports the use of a text box within the pre-scripted options. It can, for example, be used to fill in a name or a number like the amount of calories someone wants to burn. This allows for more interaction for the user and more intuitive answers.

As choosing from a set of options is the current way of interacting with the system and typing and voice recognition are not available without a large rework, choosing options has to be translated into VR. The system should thus provide the player with some options in VR and there are several ways of doing this. The current system lists the option underneath each other, each as a separate button to click on, this is very simple and mainly in place for testing purposes. Some possible alternatives include: a dropdown menu and full screen options.

Having a dropdown menu works especially well when there are many options and the page/screen becomes cluttered otherwise. As COUCH is using three options, a dropdown menu would only make the interaction slower (the users have to ‘click’ more, first open the menu, then select the option). It could also be more difficult to read the different options, as they are not instantly visible (menu has to open first) and the options might be closer together in such a menu, making it harder to distinguish and select them.

Presenting the options full screen to the user might work well on a normal screen, as the attention is briefly completely shifted towards reading and choosing an option. In VR this might introduce two problems, increasing the chance of motion sickness as the screen might suddenly change and decreasing the engagement for the same reason. Most answers are also short in their formulation, which does not seem to warrant a large visual representation and reduces cognitive load for the user. Moreover, according to Oculus, a pioneer in consumer VR headsets, released a guide with guidelines on how to make the best possible VR experience [35]. They state that developers should refrain from using a head-up display (HUD) and instead integrate UI elements in the scene. A HUD is a method of displaying UI, in which the UI elements follow the camera and appear stuck in place. This is often used for elements that should be visible at all times, such as a timer. However, this only works well in non-stereoscopic games/applications, as the HUD is easily differentiable from the background. In VR this causes a problem, because the HUD has to be visible for both eyes and thus needs to have some depth, which can cause it to collide with objects in the scene, causing discomfort for the user. This makes buttons as HUD elements not feasible.

This leaves the existing option of using separate buttons in the scene to represent the different options. This should provide the easiest way of interacting and having only three buttons on your screen, overlaying the environment, should not obstruct the user in any way. How users will interact with the UI elements will be further explored in section 3.2.

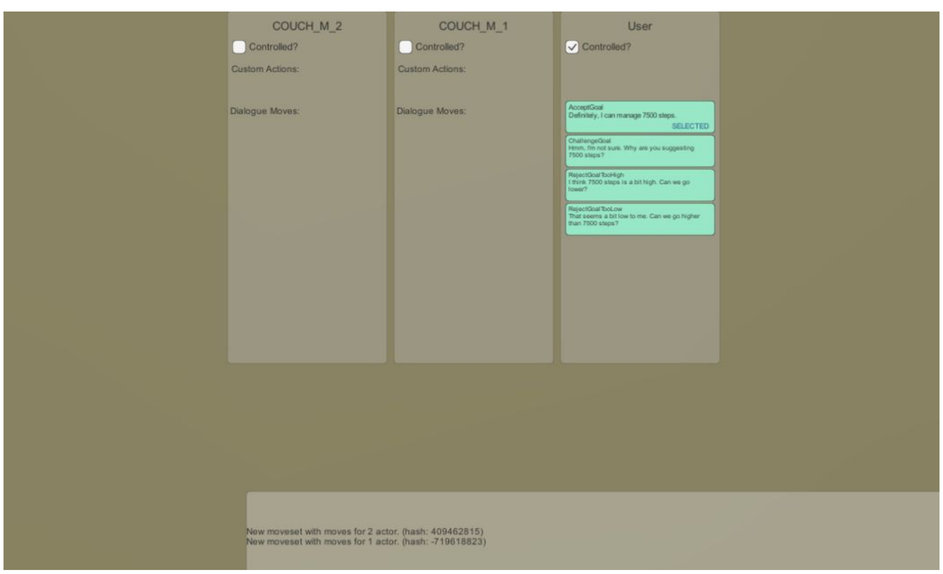


Figure 5: Current User Interface

However, while the UI already has an implementation, the Graphical User Interface (GUI) leaves much to be desired. The way the UI looks in the current version is very basic and rather complex, as can be seen in figure 5. It is designed to work well for testing and to provide as much information about the state of the system as possible. When switching to user centered design only the UI elements that are needed to use the conversation, are necessary in the system, as there should no clutter and there should only be one focus at a time, especially for older adults [36].

Buttons

The buttons that are used to represent the possible options should be simple, big and distinguishable. They should be simple as they only have to be able to show the possible answer written down (e.g. "2000 steps sounds good!") and clutter should be removed to reduce cognitive load [38]. Furthermore, they should be big to make sure the button is clearly visible and the text is clearly readable. It is widely known that eyesight deteriorates with age, which warrants the need for large buttons and fonts in COUCH. Lastly, the buttons/different options should be easily distinguishable, as the user should understand the difference between them and not select the wrong one. This means there should be some spacing between them and they should have a slight colour difference between them, however, it is important to keep the colours consistent in order to prevent confusion: a 'negative' answer could be slightly red, as long as future 'negative' answers have the same colour. This is also applicable to their location on the UI: the same category as answers should appear in the same location (e.g. left), to add consistency.

When looking at the location of the buttons, they can either be stacked horizontally or vertically (as it is now). The benefit of placing the buttons above/under each other is that it is the most space efficient. It can support more and longer options compared to placing them next to each other. However, COUCH currently only offers three options to choose from and the options are short sentences. This means the buttons can be placed horizontally, like in many other applications. As the users can turn their head (slightly) in VR, this does not cause a problem. In order to prevent certain options from 'disappearing', all buttons should be at least partly visible within the user's direct field of vision. And fully within the range in which users can comfortably move their head [38].

Another aspect to consider is the availability of the buttons. Should the buttons always be available, even when there are no options to show on said buttons or should the buttons disappear after an answer is given? To give the user time to reflect on their chosen answers and to provide more context to the advice given by the coaches, whether the last selected option should always be visible is interesting to investigate.

Subtitles

In order to make COUCH and VR accessible to everyone and adhere to the guidelines [36], it needs to be usable without having audio. This is especially important for those who suffer from hearing impairment. Therefore, the system should provide subtitles when the coaches are speaking. This makes the system usable for the hearing impaired, but it can also help other people understand the coaches better, as the user does not only hear them, but can also read along. The subtitles should not appear obstructive and should be presented above the options, so that the user can read the question while the options are on screen and no problems of occlusion would arise. The subtitles should also be clearly readable and show who is talking, a text cloud like the ones found in cartoons would best here to provide contrast and context. Another option would be to have the subtitles as a HUD element. This would mean the subtitles are always visible (even when turned around), but measures need to be set in place to prevent occlusion with the scene. In both cases, the subtitles should also clearly state who is talking by adding their name. This is however, in direct violation of the Oculus guidelines.

Previous Questions & Answers

Another (possible) UI element is the possibility to view the previous question asked and answers given. It can be useful for the users to see the conversation so far and not contradict what they said earlier. As this is something extra to aid the user and should not distract, it should not be positioned in their direct field of vision, but rather in their peripheral view. This makes it so it is not cluttering the screen but is still available within comfortable head movements. As this is another UI element, the HUD should be avoided, which means it can either be hovering in the air somewhere, or be integrated into the scene. An option for this would be to hang a poster or painting on the wall containing this information.

4.2 Interaction

One of the benefits of VR is that it allows for greater and different interactions with a system. A controller in VR allows for far more complex interactions with objects and environments, where a mouse or joystick can only provide an two dimensional input (x and y). A VR controller can not only provide its position in three dimensions (x, y and z), but can also provide three dimensional rotation. Additionally, VR controllers are easy to use and intuitive due to their natural, direct mapping to hand motion[37]. This is taken into account when ideating how users are going to interact with the system. Currently, the system requires the users to click on their preferred button with their mouse. As a mouse does not work well in VR, this needs to change.

Most commercial VR headsets come with controllers, this means COUCH can utilize controllers and still support many headsets and for this thesis, an Oculus Rift S is available, which comes with two controllers. There is a way to work around controllers and that is to use ‘gaze’ as a way to select UI elements. This would mean the user sees a timer or icon appearing when they look at a button/option (in the case of COUCH) and when they gaze long enough it ‘presses’ this button. Using this technique could be useful in certain scenarios, but it could also trigger accidental ‘presses’ and be difficult to use for people who have difficulty sitting still or have a lack of focus.

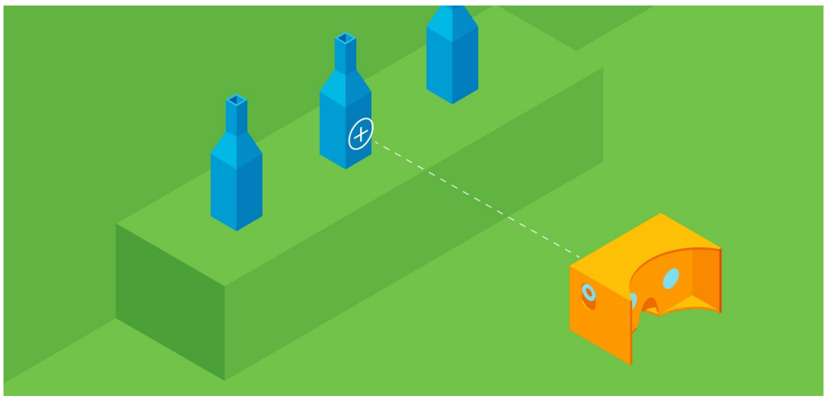


Figure 6: The ‘gaze’ functionality in VR; source: Google

The controllers that come with VR headsets can be used in a variety of ways. The ‘gaze’ discussed above can also be implemented using a controller, commonly referred to as a pointer. Many current games and applications use this feature to mostly navigate the various (complicated) menus. It is quite an intuitive solution and a good supplement for a mouse. However, it works best for large menus which would otherwise be complicated for the user to interact with and it might not work that well for older adults, as hand eye coordination and head steadiness decreases, which are both important for this feature to work.

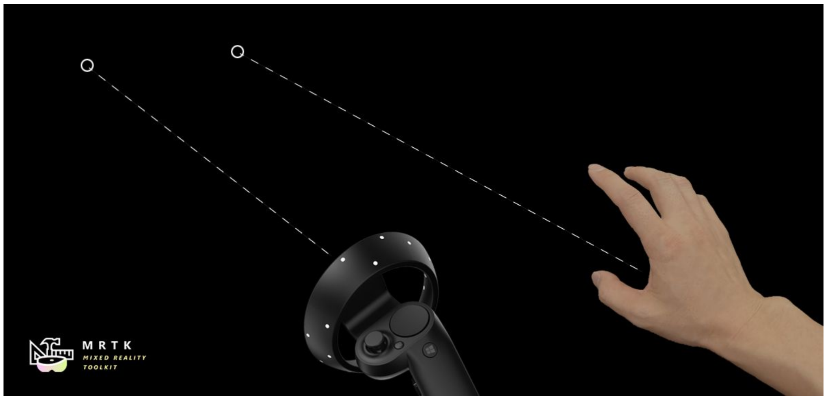


Figure 7: A 'pointer' in VR; source: Microsoft

Many different types of interaction are possible with these controllers, as they have many features themselves. The controllers of the Oculus Rift S, like many others, have triggers, buttons, a joystick and touchpad, offering many possibilities. Different buttons could be mapped to certain options, the joystick could be used to scroll through the options, etc. These features are, however, designed to work in addition to the 3D interaction of the position and rotation of the controllers and they have no added benefit in VR over conventional display and interaction technologies. Therefore, the 'simple' interaction COUCH needs to select buttons should use the 3D location of the controllers.

When using the controllers the most used interaction is to 'grab' the desired objects. In the case of COUCH, this would enable the user to grab the option that best suits them. Many applications use grab to pick up objects and/or hold them, so the question becomes whether this would work well for COUCH, as the user would not be able to pick the options up, but instead just select them (change colour e.g.). Grab is an intuitive solution for many kinds of interactions in VR and is interesting to see if this translates well into picking an option. The interaction should feel natural and logical to the user, even though they might not pick up the option if it were an actual physical button.

Instead of grabbing the buttons, which might feel unnatural or illogical, one could also 'push' the buttons. This means the option would be selected when the user touches the button with their virtual controller, which might be a more natural response to a physical button also (trying to push or press the button). It is preferable to have the interaction that requires the least amount of explanation and the easiest learning curve.

4.3 Visualization

Another way in which the current implementation of COUCH can be changed is its appearance or the visualization of its various elements, such as the buttons and subtitles briefly mentioned in section 3.1. The current system is built as a prototype and the visualization is therefore very simplistic and unintriguing. In an effort to make the system more appealing and engaging, this will be changed. The buttons mentioned in section 3.1 should be a simple shape that does not distract from the content written on it, likewise the colour should not be distracting, but provide enough contrast to make the content readable. The other areas that need attention include the location of the scene (or the background), the users themselves and the controllers.

Location

The location the current conversations in COUCH take place is quite simple and could be unnerving to those who are afraid of large open spaces (a symptom of agoraphobia). It is also cold in its colours, with only the other coaches providing some colour into the scene (see figure 5). This part of COUCH is the main subject of the bachelor thesis of fellow student Timo Petersen and will therefore remain quite simple in this thesis. The setting of the scene can take many forms such as a doctor's office, a beach, an empty room, ect. Most important to this research is to understand what changes affect the user engagement and the difference between the possible options might be hard to quantify, meaning only a couple choices will be considered.

Having a clean room, like the current one, makes the player focus on what is really important: the conversation. It removes all visual clutter, but it is currently counteracted by the in-game browser and the obtrusive UI. While it does provide the user with a central focus, which is especially useful for older adults [36], but can also cause issues as mentioned above. In order to solve this, the current room should be replaced for something a bit more enclosed and grounded. The user should not feel trapped, which can be the other extreme of agoraphobia, meaning the scene should be in some kind of room, with a clear view of the outside world. Therefore the room should be simple, with basic shapes and cool and discrete colours and include a window with a view.

Given the problems of simple and cold environments, another option can be proposed. In order to make the user feel at ease, a warm, inviting room can be used. The challenge is to make it cosy without it distracting from the conversation too much. For this, a cosy room with mostly wood, a desk and bookshelves might fit perfectly. The colours of the wood and books add to the cosiness and the desk implies the room is still a professional/working area.

The other features were divided over these two locations where they compliment each other and try to minimize the cognitive load on the users. This is further explained in section 4.5.

The Hands

Another interesting feature to consider is the VR visualisation of the user's controllers. There are many different shapes that can take its place in VR, but the two most used ones are a 3D model of the same controllers the person is using, or hands that animate when you press/touch certain buttons. Both have their advantages and disadvantages, the controllers will give you a very one to one feeling as you are holding the same thing in VR as in the real, it can however feel less natural than seeing hands in front of you. Hands have the advantage of being more realistic and fun, but can also cause disconnect as your hands in VR might not reflect what your hands are physically doing.

4.5 Concepts

From the large amount of possible solutions explained in this chapter and the lack of an obvious preference for many of the categories, it became clear that one version might not be able to represent the best options available, without extra user testing. Two concepts however, would be able to entail most of the solutions proposed and would create a good method of comparing different features while still being able to answer the main research question of what the effect of VR is on user engagement. The two concepts were formulated and the different solutions divided between the two. A quick sketch of both concepts was made to add clarity to the design process, concept 1 illustrated in figure 8 and concept 2 illustrated in figure 9. A list of all the features of each concept can be found below in figure 10.

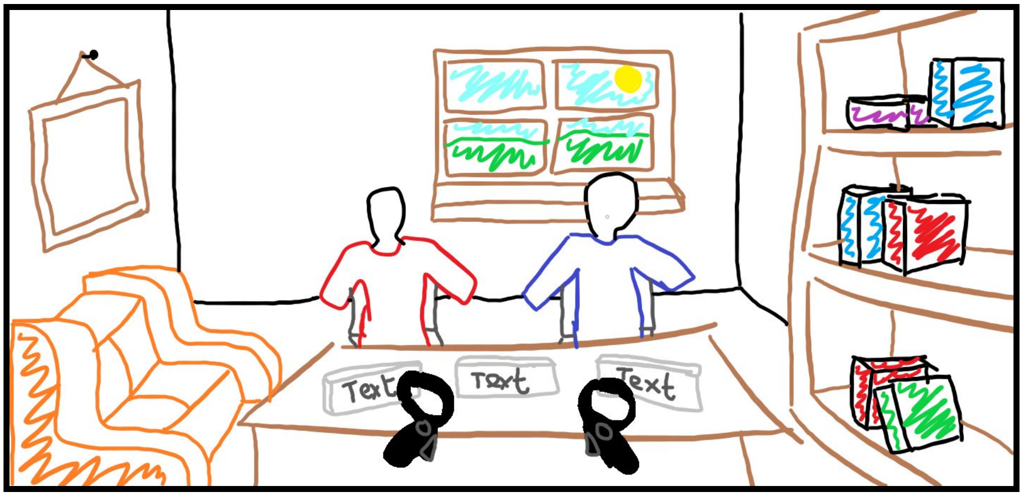


Figure 8: Illustration of concept 1

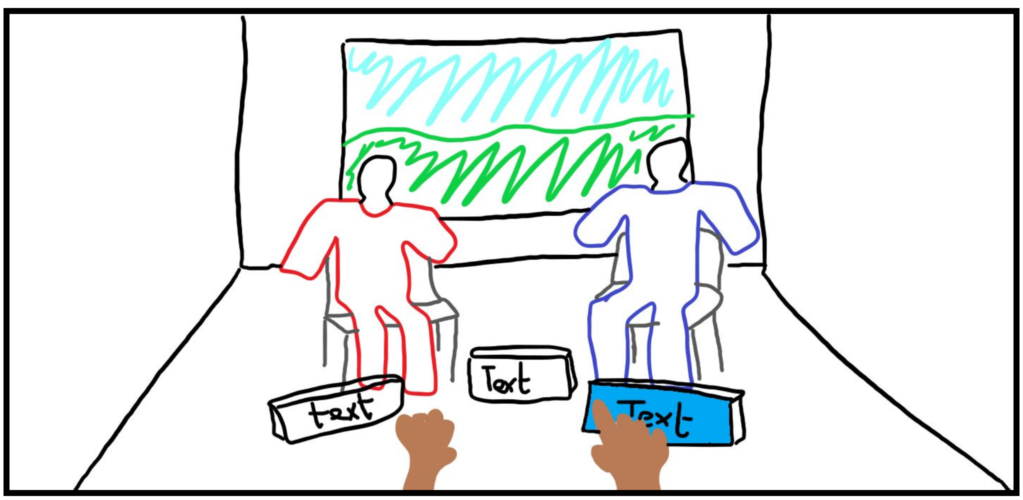


Figure 9: Illustration of concept 2

	Concept 1	Concept 2
Button Location Buttons are displayed in 3D and will be placed in a row next to each other.	**On a table** in front of the user.	Hovering **in the air** where they do not obstruct the view of the coaches.
Button Timing Whether the buttons are visible at all times or only when the user is prompted to answer.	**Always visible**, even when no questions are being asked.	**Only when prompted** and disappear when the question is answered.
Previous Question & Answers Whether the previously asked questions and given answers are being displayed somewhere on screen.	**Yes**, the previous questions & answers are displayed.	**No**, the previous questions and answers are not displayed.
Subtitles The display mode of the subtitles of the conversation.	**No** subtitles, can use previous questions and answers.	Placed in **worldspace** where it is not obstructive.
Interaction Mode Controllers are used to interact with the buttons.	**Grab** the buttons to select them.	**Push** the buttons to select them.
Location The location of the scene in which the conversation takes place.	A **cosy work room** that is warm in colour makes the users more at ease.	An **empty/clean room** that minimizes distractions.
Controllers appearance The controllers can appear as many different shapes in VR.	Appear **as controllers**, the same ones in VR as in real life.	Appear **as hands** instead of controllers.

Figure 10: Table of features and their realisation

Some of the features that might work well in conjunction with each other have been added to the same concept. Concept 1 has the buttons be placed on a table in front of the user so that they are easily reachable to grab and will display the controllers as controllers, as the goal is to move them inside the buttons. With the buttons on a table/desk, a cosy room fits in well and letting the button remain on the table after selection could be considered more natural than in floating in the air. It also might be less distracting in the cosy room as the button jumps out less. For concept 2, the clean room is chosen where the floating buttons make more sense. To push the buttons is also a good combination with showing the controllers as hands, as it feels natural to push something with your hands or fingers. The subtitles and the backlog (previous Q&A) are two features that resemble each other a lot and should be separated to prevent one from overshadowing the other. Having the backlog in concept 1 makes more sense as it fits perfectly in the painting frame and can be off to the side where it is not distracting. The subtitles fit in concept 2 well as the room is empty and thus has space for a text cloud.

Chapter 5: Specification

To prepare for the realisation of the concepts into prototypes, requirements should be established to ensure a smooth process. This section exists of software, hardware and user requirements.

5.1 COUCH

At the beginning of this thesis, the newest demonstrator version of the COUCH system was provided to the researcher using GitHub. During the thesis the system was upgraded twice, but the core functionality and setup remained unchanged. COUCH is a rather complex system as it uses many different modules that each have a specific purpose and need to communicate with each other. Only the modules that have a direct impact on this project will be explained in simple terms, to provide extra context, especially in the realisation phase.

In order to run the system for parts need to be running (in order):

- DAF: Dialog and Argumentation Framework, which stores the content of the dialog and allows for complex argumentation responses from the coaches.
- ASAP: Controls the ASAP agents (both of the coaches in this build), by moving them and letting them speak.
- Unity: Handles the interaction between the system and the user.
- Flipper: Handles the information between all systems.

5.2 Autodesk Maya

Maya is a 3D modeling application from Autodesk. It is free with a student licence and used by professionals in, among other things, the gaming and film industry. Maya will be used to create all the assets that fill populated the scenes of both prototypes. These assets will be sent to unity as whitebox assets (without textures or materials). A list of assets per concept can be established:

- Concept 1 (cosy work room)
    - Simple button
    - Desk
    - Book
    - Bookshelves
    - Lamp

    - Plant
    - Couch
    - Painting
    - Window frame (old)
- Concept 2 (empty/clean room)
    - Clickable button
    - Window frame (modern)
    - Lamp

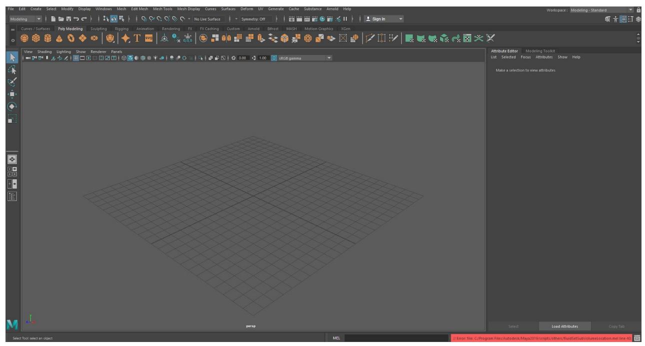


Figure 11: User interface of Maya

5.3 Unity3D

Unity3D (or Unity) is a free game engine that focuses on versatility. This means developers have a lot of freedom and Unity supports the newest technologies, such as VR. The COUCH demonstrator already has an existing project running on Unity version 2017.4, this version of Unity and the existing project will thus be used for this thesis. A list of features that need to be implemented using Unity can be established:

- Main features
    - Implement VR looking and walking around (Oculus plugin)
    - Make the options appear on the buttons and scale buttons accordingly
    - Import Maya models and add materials to them
    - Create two scenes and use lighting and composition to make them look good
- Concept 1
    - Have the text from the user and the coaches display properly on the painting
    - Display the controllers as controllers (Oculus plugin) and add trigger functionality
    - Make the button highlight when selecting them
- Concept 2

- Have the text from the coaches display properly on the text cloud (subtitles)
- Display the controllers as hands (Oculus plugin) and add collider to index finger
- Make the button respond to pushes

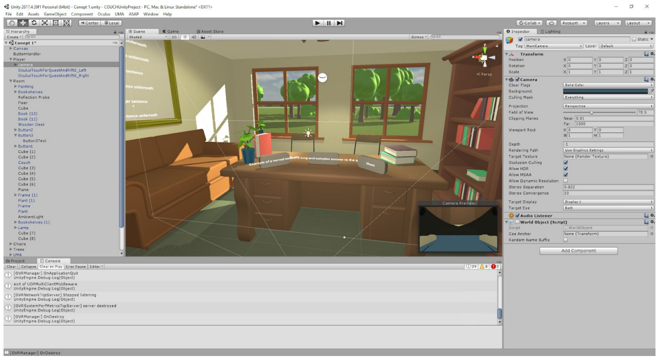


Figure 12: Couch project in Unity

5.4 Adobe Captivate

Adobe Captivate is software by Adobe that allows for interactive presentations and, in this case, interactive video. Captivate works with slides, meaning each slide could play a video of one of the dialog options. While not being the most suitable for the needs of this project, Captivate gets the job done and will work reasonably well.

5.5 Oculus Rift S

The Oculus Rift S is a Virtual Reality headset and the successor to the very popular and innovative Oculus Rift. It is a well rounded pc-tethered VR headset that utilises inside-out tracking, meaning it uses its cameras and sensors to calculate position and rotation (6DoF). It also comes with two controllers with many interaction possibilities.

Figure 13: The Oculus Rift S with controllers; source: Oculus

Chapter 6: Realization

This chapter will go over the process of turning the concepts, proposed in chapter 3 and specified in chapter 4, into prototypes to be used in the evaluation. It will also explain the different modules created and how they relate to each other.

6.1 Procedure

This section will go through the process in a chronological order, but some features may have been added before or after the proposed order.

Creating the assets

The list of assets formed in section 5.1, were created first, as some of the design decisions were not yet taken at this point. The assets were all created in Maya and were all produced in roughly the same style: not too realistic. Games or scenes are often considered ‘bad looking’ when it does not quite meet its intended goal. For this reason, the models were intentionally imperfect and the shapes simple. A few examples are shown below (still in whitebox).

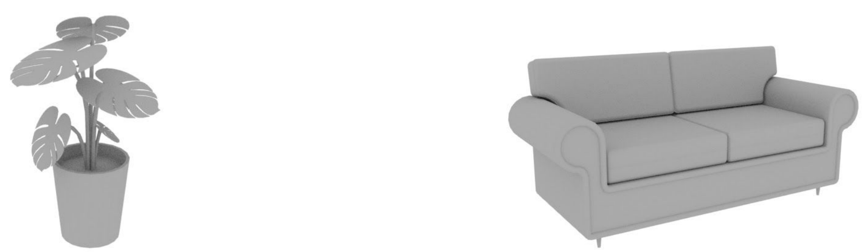


Figure 14: Whitebox models

After this all models were imported into Unity where they were given a material. For the same reason as described before, the material was kept very simple, the only difference between all materials are the colour and the reflectiveness. This made sure the models and their textures always look like they belong together. Below you can see the same models, with materials in Unity.

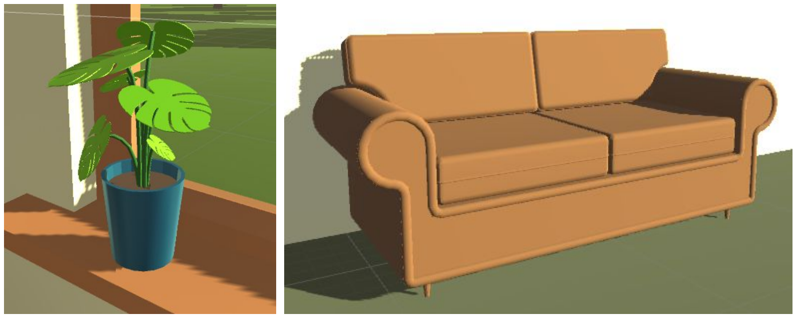


Figure 14: Models in Unity with materials

After all textures were imported, basic layouts of the prototypes were formed and all unnecessary elements from the old scene were removed, such as the internet browser.

VR integration

Getting VR working in Unity is actually quite simple, especially when using Oculus headsets. First, the ‘Oculus Integration’ package was installed, which gives access to many great prefabs to use, but also adds oculus headsets as an option in Unity’s settings. Once enabling this setting and removing a problem that was caused by the import, the camera in the Unity scene responded perfectly to the movement of the headset. The camera however, was not in the correct position and this was solved by setting the anchor point on the floor. The default scaling was already good: the researcher was slightly taller than the coaches.

After this, controller model prefabs were imported into the scene (included in the Oculus Integrator package) and a script was created that displayed them with correct position and rotation. Another option was to use the controller prefabs that already came with the controllers, but as this project will use very simple controls, simple scripts can provide clarity and make it easier to debug if something goes wrong.

Next, although only for prototype 1, the controller received a ‘grab’ script that communicates what objects they are interacting with, such as buttons (and which button). In order to select certain options, the trigger input of the controller was also gained through scripting. Now the program knew which button the user was selecting and whether the trigger was pulled.

Building the prototypes

With the controllers working as needed (for prototype 1) work continued on adding functionality, a script was made to handle the behavior of the buttons, that was later altered to also work with prototype 2. The text that was being sent to the UI would now display on the buttons and buttons could now be selected and would trigger the right response, meaning the prototype was functional from this point onward.

It became apparent at this stage that the text on the buttons would not disappear behind other objects, such as the controllers. Sadly, the version that Unity was running on did not support the newer package manager, without which this problem could not easily be solved. Instead, a custom shader was created that made sure the text was only rendered when in front of other objects.

Next, the buttons were made to disappear when no questions were being asked, except for the chosen answer in prototype 1, and the controllers in prototype 2 were changed to hands, also a prefab from Unity Integration. Support for subtitles and a backlog were added to the button handler. At this point the shader was modified to support coloured text, which was needed for the backlog.

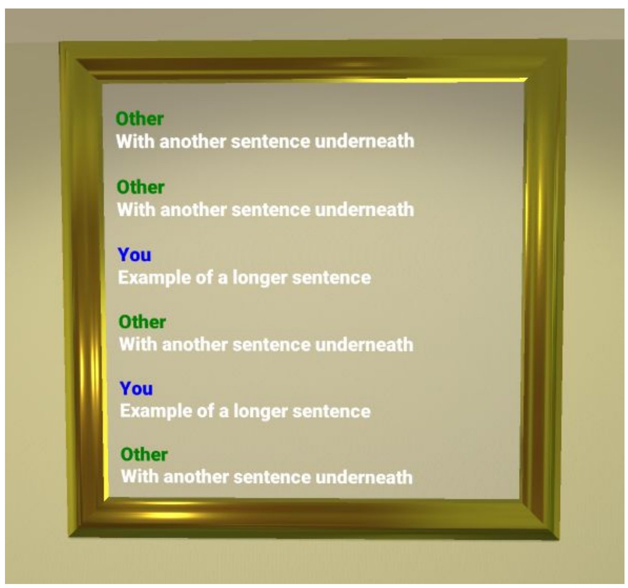


Figure 15: The painting with the backlog in Unity

So far, the buttons remained the same size no matter the length of the sentence. A template script was created which would cut up sentences if they became too long and sized the button accordingly. This script was then also used to scale the subtitles correctly and later (after the pilot test) to scale the backlog properly.

Pushable buttons

The last thing added was the interaction of prototype 2 and this was one of the harder features. The goal was to have a very tactile button that one could press, similar to a physical button. To achieve this, first the hands needed to get a collider attached to the index finger, with which the user can press the button. In order to make the button feel real, the button should want to push back and return to its original state, as if it had a spring inside it. For this reason the button was given the configurable joint component, which acts like a spring when set up correctly and allows for much customization. After some tweaking and testing the spring worked as intended, only moving over the correct axis and not surpassing its boundaries. The finished version can be seen below.

Figure 16: The effect of pushing a button and it bouncing back

Lastly, the prototypes were cleaned up, issues were fixed and they received a graphical overhaul, by playing with light sources until the preferred look was achieved. The prototypes could have looked better if post processing was used, but sadly this was not available with Unity missing the package manager. Comparing the first early concept art to the finished result in the figure below, it is clear to see that the overall goals were achieved.

Figure 17: From concept to prototype, comparison

Dialog

To prepare for the testing phase another thing became apparent: the current dialog was too short and too simple, only lasting about thirty seconds. In order to keep the users' attention for longer during the tests while avoiding to reset the dialog over and over, a custom, longer dialog was created. With some major hiccups that caused delays of the user tests, the dialog was finally ready with working variables (which was not possible before), which also meant the interactive video could be created. The figure below shows the dialog diagram, a larger version can be found in appendix A2.

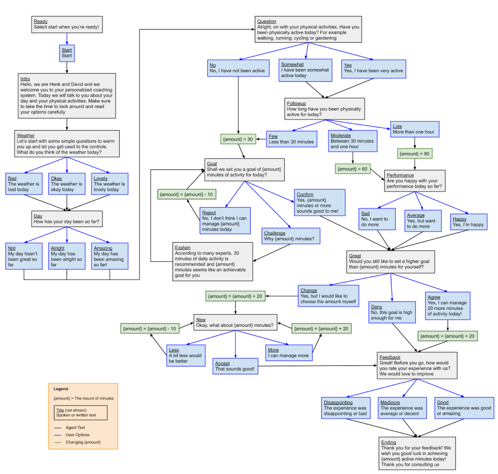


Figure 18: The final dialog

Interactive videos

In order to show the VR installation as best as possible to those who cannot be approached using VR, interactive videos are created using Adobe Captivate. To make this work, first all scenes and possible choices were recorded for both versions, after which they were cut and linked to slides in Captivate. These slides were connected to each other and buttons were added to jump to specific slides. Sadly no audio could be recorded while using the VR headset, so the audio was recorded separately, and synced up with the existing footage (and the right person speaking). This process took a long time, but does provide a more engaging and interactive experience for the users compared to just watching a pre-recorded video. By continuously asking for input, the user has to keep paying attention and play a more active role.

Figure 19: Snapshot of the interactive video

6.2 System overview

An overview was created that shows how the different scripts communicate and how each script relates to the others. This overview can be seen below. Legend:

- Red - COUCH system components
- Blue - Shared components between prototypes
- Green - Prototype 1 scripts
- Pink - Prototype 2 scripts

Button handler is the main script that handles all the information passed between the different scripts. It pulls data such as possible moves from the UIMiddleWare script and sends data like the next move. Button scaler, backlog and subtitles all process text they receive from button handler and put it in the proper place and format. The controllers and grab script together sent the buttonhandler which button is currently being pressed.

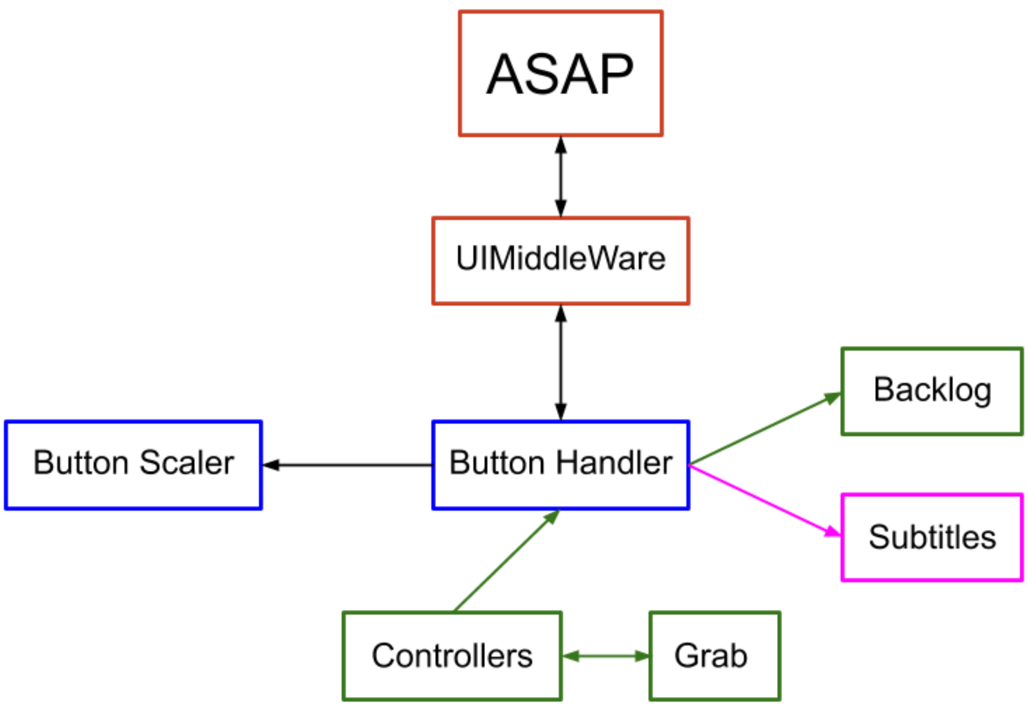


Figure 20: Prototype script diagram

Chapter 7: Evaluation

In the evaluation phase, user tests will be performed in order to see the effect of the produced prototype on the users and answer the research question proposed in chapter 1: *"What is the effect on user engagement when implementing VR in the Council of Coaches system?"*. This means the goal of the user test is to find out the user engagement of both prototypes and what features have an effect on this. In an ideal scenario the prototype would be tested with the target group (older adults) directly, as this would provide the most valuable feedback on the system. However, because of the current COVID-19 pandemic, this is not possible without attaching health risks. This means the prototype has to be tested with other users than the target group and are in line with the COVID-19 guidelines: younger adults as proxy users. These 7 users (including pilot), although not many, live in the same house as the researcher and can therefore be used as participants in the test, whilst complying to the guidelines.

While proxy users are a great way to simulate the target users and can provide valuable feedback and impressions on the prototype, they are always (somewhat) detached from the actual target group, as their ages range from 19 to 23. Even though the participants will be told to act as if they were an older adult (as will be explained later), their preferences might not be representative of the target group. In order to get the opinions of older adults, another test will be performed, this time using 3 online video interviews (including pilot), showing the participants interactive videos. This allows the researcher to contact the target group while adhering to the COVID-19 guidelines. While not being to go as in-depth as face to face user tests on the VR part of the prototype (as the users cannot use a VR headset and have to view the prototype on a simple screen), it can still provide valuable feedback for further iterations of the prototype and can help answer the research question.

In order to have the user tests be executed properly and consistently, the setup and procedure of both tests will be explained thoroughly.

7.1 Methodology

The setup and materials, as well as the protocol of the evaluation procedure can be established for both tests in such a way that it is easily replicable to add consistency.

7.1.1 Setup and Materials

Both the user tests require different materials and setup, but share many similarities. The goal is to have the two experiments resemble each other as closely as possible so that the results could be compared between the two, thus increasing the sample size and credibility of the conclusion of the evaluation. This can only be done to certain sections of the evaluation if the results of both tests show clear similarities.

Proxy user test

In short, this test will have the participants (proxy users) use the two concepts created in VR and give their opinions on both. A list of materials can be established:

- Computer that runs the COUCH system with the concepts
- VR headset connected to the computer
- VR controllers and extra batteries
- Chair for the participants to sit on
- Device to record interview audio
- Informed consent form (and pen)
- Interview questions
- Reward for participants (candy bar)

The user test should take place in a(n) (empty) room where there is enough free space for a VR setup, around 2 by 2 meters. Even though the participant will be placed in a chair and the system is designed to work best from a seated position, they are not told to remain seated and thus need the space to walk around if they please, without hitting obstacles. If this occurs, the researcher will make sure to assist when needed and prevent participants from falling or bumping into obstacles. The chair should be placed in the middle of the room and the researcher should calibrate the room beforehand to make sure the participants will be seated in the correct location of the VR scene as well. The researcher themselves, along with the computer, should be seated far enough away from the participant to not obstruct their movement. All other materials mentioned should be present and working properly.

Figure 21: Proxy user test setup, participants sits in the chair on the right

Online target group interviews

The online interview will be a more simple version of the proxy user test and will mostly focus on the interview. An interactive video of both the working concepts will be recorded beforehand and will be shown to the target group instead of the VR experience. This interactive video will walk the participants through a conversation, just as the proxy users do in VR. While doing this, the videos try to cover and show as much of the features as possible and make the system understandable for the older adults (e.g. no sudden head movements, clear/slow interaction, long attention to features, etc.). To make the videos more engaging, the videos will be interactive by letting the participants choose the answers. This is important as, for VR especially, a video can never truly convey the experience of a (VR) game or application, which might be even more difficult for older adults. The materials needed are:

- Computer with webcam
- Video call software (as agreed upon with the participant and within UT guidelines)
- Interactive videos of both concepts ready to screen share
- Video/audio recording software to record the interview
- Informed consent form (online)
- Interview questions

7.1.2 Recruiting

Important for the user tests is that the right people are chosen to participate in them. Due to the current COVID-19 pandemic, this process is severely more complicated and might therefore differ somewhat from the desired state. Luckily, due to the nature of this research, social bias should be minimal, as both concepts and all its features are made by the researcher and the goal is to see which version and features work best in terms of user engagement. This means housemates of the researcher can be used as reliable participants and every person within that group that is willing to participate is accepted.

For the online target group interviews some recruiting is required. However, as bias should not play a role, the participants can be relatives of the researcher or other acquaintances. The only requirement for these participants is that they are older adults (50+), that they are in good enough health to conduct an online interview and have sufficient understanding of technology to engage in a video call. These participants can be approached by phone.

7.1.3 Protocol

The procedure of the experiments will mostly be the same between the two tests, with some difference in how certain actions are performed. This procedure can be divided into multiple steps (in order) which explain in more detail what both tests will look like.

- Introduction
- Informed consent form
- First short instruction
- First concept testing
- First short questionnaire
- Second short introduction
- Second concept testing
- Second short questionnaire
- Semi-structured interview
- Debriefing

Introduction

When the user tests start, the participants will receive an introduction shortly explaining what the test will be about and what the research aims to achieve. They will also be given an overview of the procedure of the test, so that they know what to expect. The introduction will be the same for both tests and can be found in appendix B3.

Informed consent form

After the introduction the participant will be provided with the informed consent form and asked to sign it, a physical paper and pen for the proxy users and an online form for the online target group interviews. This consent form includes everything the participants need to be aware of and allow the researcher to use the anonymous data gathered to draw conclusions. The informed consent form can be found in appendix B2. An information brochure is also provided for the participants to read and take home, this can be found in appendix B1.

First short instruction

Next, for the proxy users, a short introduction is given that explains the basics of using the VR headset and how the participants can interact with the prototype. For the online target group interviews, the mechanical side of the prototype (how to interact in VR) is not explained, but rather the features of the prototype are explained in more detail, as they cannot interact with them to understand them better. The written instructions for both concepts can be found in appendix B4.

First concept testing

For both tests the participants get to see/interact with the first concept prototype: the proxy users get the concept prototype loaded onto the VR headset and are allowed to interact with it for the duration of the given dialog (a couple of minutes), while the online participants are shown an interactive video of the concept. In order to reduce the possible learning effect that could arise when switching from one concept to the other, half of the participant pool (for both tests) should get concept 1 first, while the other half should get concept 2 first, also known as counterbalancing. The participants are divided between the two starting concepts by having the first participant start with concept 1, the second with concept 2, the third with concept 1, etc. During the handson with the prototype the researcher will observe the participant and note any interesting observations.

First short questionnaire

To get the initial impressions of the participants, a short questionnaire on user engagement taking a couple minutes maximum is conducted in which the participants rate the experience on multiple subjects. The participants will be asked to fill in the questionnaire on a (by the researcher) provided laptop. This short questionnaire was created and tested by O'Brien et al. [39] and can be found in appendix B5.

Second short introduction

An introduction given for the second concept prototype, only explaining the new features/interactions.

Second concept testing

For both tests the participants now get to see/interact with the second concept prototype for a couple minutes, which is the prototype they did not get to see/use in the first test round. Observations will also take place.

Second short questionnaire

The same questionnaire of the first concept will be conducted in the same way, this time focussing on the second concept prototype.

Semi-structured interview

After both concepts and questionnaires are completed a semi-structured interview will be conducted. In order to refresh the participants' memory, a picture of both concepts will be provided during the whole interview, which can be used as reference. This interview focuses on the comparison between the two concepts and asks the participants what they think of the features (as outlined in figure 5) and which feature realisation they find better and why. Concludingly, it will also ask which concept overall they like more and why. The interview will focus more on the interaction for the proxy users and more on the visual appearance for the target group as those questions will be easiest to answer for each group. Questions used in the interview can be found in appendix B6.

Debriefing

At the end, the users are thanked for their cooperation and will be reminded what will happen with their data collected and how they can get contact for further questions. The proxy users will also receive a small compensation for their time in the form of a candy bar.

7.2 Execution

7.2.1 Pilot tests

Before testing with the intended users, a pilot test will be conducted for both tests in order to assess the testing procedure. The feedback from the pilot will be used to improve the testing procedure and fix possible issues. It is also a good way to practice for the actual tests and ensures they proceed smoother as a result. These pilot tests will be conducted using a single user, who was selected using convenience sampling and is not part of the participant group.

Proxy user pilot

The pilot test of the proxy user user test was executed using the same procedure as the normal test, but by adding the ‘thinking out loud’ technique to get feedback on the procedure. Furthermore, the interview was extended to allow for extra feedback on the procedure. These remarks were noted by the researcher.

Overall, the results from the pilot test were very minor, with only a couple of remarks. The first is that the procedure could have gone a lot smoother if the researcher had a better overview of the procedure and what text to say when. This overview was created for the actual tests, aiding the researcher and providing more structure and consistency to the tests. The texts that were used (that can be found in the appendix as mentioned before) were all written in English and the participants were all Dutch. This added complexity as forcing the tests in English might influence the answers given and translating these texts to English on the spot might cause inconsistencies. Therefore, all the texts were translated to Dutch and added to the overview. The questionnaires were kept in English, as most participants understood all questions and not translating it would ensure to preserve the integrity of the questionnaire.

Secondly, some confusion occurred when calling the two prototypes 1 and 2. This was especially confusing as the pilot test started with prototype 2 (as determined by a coin flip), meaning the words ‘one’, ‘two’, ‘first’ and ‘second’ became confusing and required more thinking than necessary. The pilot tester mentioned a possible solution, which is to call the prototypes A and B. This way, no more confusion should arise when discussing the prototypes. This solution was then implemented.

Besides the slight tweaking of some dialog options, the procedure required no further changes according to the pilot test, but some technical issues were found that needed to be fixed before the actual tests. The first issue appeared to be an oversight in prototype 1, where the backlog (painting) showing the previous questions and answers was not properly tested with very long sentences, as they would fall off the painting and not be visible. As the backlog uses a slightly different way of displaying text than the buttons for example, code could not simply be copied, but had to be reconstructed and altered to make text fit the painting frame.

The second and final issue was harder to pinpoint; the dialog would sometimes stop after selecting a button. After analyses and multiple tests to recreate the problem, no clear causation could be found. Pressing the 'lower goal' button too often would always crash the dialog, but other buttons could sometimes also crash it. Sadly, without a proper causation and with a complex system, no solution was found. In an attempt to minimize the crashes, participants would be told after pressing a 'lower goal' button to not press it again.

Online target group pilot

The pilot for the online target group interviews was performed a few days after the proxy user tests and thus after the changes of the proxy users pilot test were integrated. The interactive video used in the pilot also included these changes. This pilot test could be conducted a lot quicker, as the procedure was the same as for the proxy user tests. The only difference between the two, is that the online target group interviews use interactive videos instead of the VR prototypes, which also means less instructions are needed before each video.

The pilot test found no errors in the procedure, but it did raise concerns about the ability of the interactive videos to simulate the VR prototypes, questioning whether can can be used as a substitute. This means future results from this test should be thoroughly analysed to determine whether they can be used to draw conclusions on the prototypes.

Another finding during the pilot is that observations are almost impossible to make when the participant has so little control over the prototype. Anyone can click a button with a mouse, meaning there is no learning curve and the interaction is kept to a minimum. Because of this, observations will not be actively noted, with the exception when something unexpected happens, meaning the researcher will still have to pay close attention.

7.3.2 User Tests

The day after each pilot test, the actual tests took place. The participants were divided into one hour time slots which allowed for the test, saving of data, preparing for the next test and a small break.

Proxy users

The proxy user tests were performed during the span of one day, using a sample size of 6 participants, with each test taking about 30 minutes all together. The test proceeded smoothly with the exception of a couple things; the dialog crashing and coaches staring at each other. For some participants it crashed once (during one of the prototypes), for some it crashed twice (once in both prototypes) and for others it did not crash at all. The dialog should not have a big effect on participant response, as it is considered an independent variable (the same in both prototypes), but it crashing should be taken into account that this might have an effect on the results.

The other problem that occurred for most participants, is the coaches not always looking at the user when speaking to them. The coaches of COUCH have a feature which lets them change where they are looking, but this was not functioning properly; when the coaches started looking at each other, they would not look back at the user for the remainder of the dialog. Quickly fixing it did not meet the desired results and thus the ‘bug’ stayed. While it occurred for most participants, only a couple made a remark about it either during testing or during the interview. This also should not affect the results too much, as the coaches are again an independent variable, but it should be taken into consideration when analysing.

In the proxy user test observations were noted. However, with some exceptions, there was not much to observe and thus the data gathered from the observations is shallow and should only be used to reinforce points made in interviews.

Online target group

After the pilot of the online tests, not much had to be changed besides a couple lines in the introduction and instructions. The tests could thus take place quickly after the pilot test. While conducting the test the warnings from the pilot about the interactive videos not properly representing their VR counterpart were kept in mind.

From the data of 2 user tests and direct remarks from the participants involved, it could be concluded that the data did not prove valuable enough to conduct more tests of this kind. This will be explained in more detail in the results section of the online target group interviews.

7.3 Results

The data gathered from the tests per person includes: quantitative data from two questionnaires (one for each prototype), qualitative data from the interview and (possibly) qualitative data from the observations. With a sample size of 10, not enough data is gathered to allow for a statistical analysis and thus a systematic qualitative analysis should be used. This section will report on this analysis and will visualise the results.

7.3.1 Questionnaires

The questionnaire created by O'Brien et al. [39] aims to capture the engagement of the participants, by asking questions divided over four categories:

- FA: Focussed Attention, how absorbed/focussed the participant was
- PU: Perceived Usability, a negative effect caused by interaction or amount of control
- AE: Aesthetic appeal, the attractiveness of the prototype
- RW: Reward factor, the overall success and longevity

The answers to each question in a sub-group can be added up to calculate a 'score' for that group and dividing it by the number of questions (three per group). An overall score can also be calculated by doing the same thing, but for all questions (divided by twelve).

Although no statistical tests could provide any reliable data, the scores can still be calculated and used to draw conclusions. The table below shows the average overall score per prototype. The users were asked to answer each question with: Strongly disagree, Disagree, Neither agree nor disagree, Agree or Strongly agree, with 1-5 score per answer, respectively. Note that the score of the PU questions is reversed, as those questions asked about the negative aspects of the prototypes and as instructed by O'Brien et al. [39].

The table below contains some averages of the full questionnaire data outlined in appendix C2. It compresses the data by dividing it over four groups, both prototypes times both tests. Each group has an average score of all four categories of questions (each consisting of three questions) and a total average score of all questions. As mentioned before the scores range from 1 to 5, with 1 being the lowest and 5 being the highest score, the higher the better.

Type	FA	PU	AE	RW	**All**
Proxy - Prototype 1	3.6	4.1	4.1	3.7	**3.9**
Proxy - Prototype 2	3.3	4.2	3.4	3.4	**3.6**
Online - Prototype 1	3.3	3.7	4.3	3.8	**3.8**
Online - Prototype 2	3.2	3.7	3.0	3.5	**3.3**

Figure 22: Result of questionnaire summarised

To put this more in perspective a bar graph was visualized that shows the different scores of the question categories and overall between the different combinations of user groups and prototypes. It is placed in this order to clearly show the difference between the different prototypes (as this is the research goal), while also allowing to compare the results of both user tests.

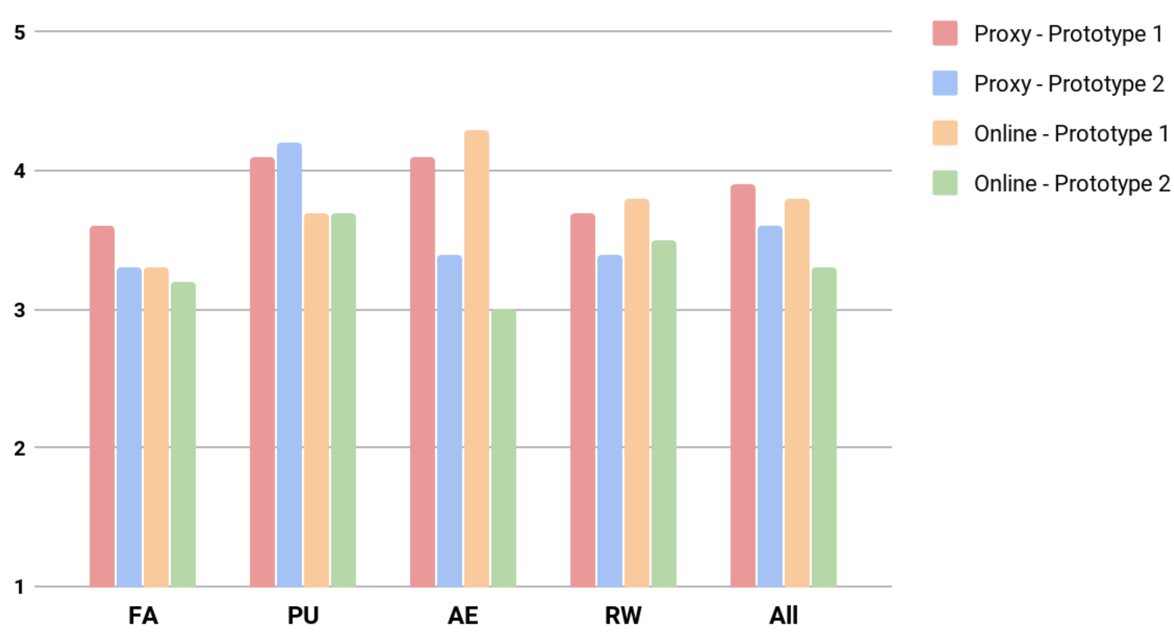


Figure 23: Questionnaire results visualised

A few interesting observations can be made when inspecting this graph. First of all the overall scores are rather high, with every category scoring higher than three points and the proxy users giving slightly higher scores overall than the online users on both prototypes. Another interesting observation is that prototype 1 scores slightly higher than prototype 2 for both user groups, which is most heavily affected by the significantly higher scores in AE (aesthetics) than prototype 2. Furthermore, PU, or the amount of control, which is largely affected by the different interaction possibilities, is rated higher by the proxy users, for both prototypes.

7.3.2 Interviews

The interviews were conducted verbally and the audio was recorded. To start the analysis the audio first needed to be transcribed and categorized based on which questions the data gives answers to. Additional remarks outside the interview questions proposed in appendix B6, were also noted and given their own separate category. The rest of this section will summarize the answers given to all questions in order and will highlight the possible differences between the proxy users and the online users, starting with the general and more open questions.

What did you think of the coaches?

While considered an independent variable, it is still important to ask what participants think about the coaches, to see what effect it might have on the rest of the test. The response from the online group was also very similar to that of the proxy user group, allowing it to be combined for this question.

The response to the coaches was divided; half of the participants thought they looked nice and reasonably convincing, while the other half thought they looked, spoke and acted like robots. Although the response seems mixed, most participants said they would take advice from these coaches and would trust them. One thing most participants agreed upon is that the coaches not always looking at you feels strange, as was expected after the results of the pilot test. The last remark that a couple participants made is that it might be nicer to have a man and a woman as coaches and not two men. This is however something that is already present in COUCH, as can be seen in figure 1, but was not available to the researcher at the time of this thesis.

What did you think of the dialog?

Just as the coaches, the dialog is an independent variable and while not being too interesting for this research, it is still a vital part of the COUCH experience.

The response to this question from the proxy users was again mixed with some saying the dialog was fine and logical, while others said it was chatbot like and too simple. Many of the participants proposed possible changes that would make the dialog more interesting and rewarding: the ending could have been less abrupt, some extra options to cover all logical answers to a question and more variation to make it less predictable. Besides all this, the participants did think the dialog served its intended purpose.

The responses from the target group were slightly different, as they mostly wanted the dialog to last longer, give them more information and dive deeper. They would also like it if the coaches told the user more about how to achieve their goals and give advice and tips, than just what the user should set as their goal. Dialog and its content seemed to be a lot more important to the target group than the proxy users, this could be due to their age, but it is more likely caused by the interactive videos. Interaction is limited which means opportunities to explore and thus the dialog becomes the focus for the participants.

What did you think of prototype 1?

A very open question with very diverse responses, also very similar between the two tests. The overall consensus was that it was very appealing and pretty and gave a cosy feeling. Besides that some participants commented that they liked the backlog (painting), for some it was too far from their peripheral vision and required too much attention to look at, while others thought it was nice that it was placed to the side, where it was not distracting and could be accessed at any time. Other features were not mentioned, but were talked about in later questions.

What did you think of prototype 2?

Two main things became clear from the proxy user responses: the prototype looked less pretty and inviting than prototype 1 and the subtitles were universally liked. The subtitles provided extra ways of absorbing the information which made it easier to understand and worked as a useful backup if you misheard something, according to the participants. Some other remarks from the proxy users stated that the lack of a desk made the conversation feel more personal and that the buttons and hands provided were a fun way of interacting.

The target group participants did not experience this and their focus was mainly on the way it looked. Besides the fact that they thought prototype 1 looked better, they both said that the emptiness and cleanliness of the prototype 2 could provide a more focussed experience for the user, with less distractions and more emphasis on the conversation.

Which prototype did you like more?

This is a very subjective question as it aims to understand which prototype was liked/enjoyed more, which could be influenced by many different factors. The response was divided as five participants liked prototype 1 more and three liked prototype 2 more. The interesting thing to note is that the participants who liked prototype 1 more, liked it more because of its appearance and because that gave them a more happy feeling, while the participants who preferred prototype 2 did so because of the provided subtitles and its associated convenience and accessibility.

Which prototype do you think would work best?

This question was again divided, but this time leaning towards the second prototype with four participants thinking prototype 2 would work best, two thinking prototype 1 would work best and two saying it depends on the person. The reasons for these preferences were the same as in the previous question, but for older adults most participants thought accessibility was the most important aspect (the subtitles).

What is your ideal version?

For this question all responses, from both tests, were almost exactly the same, only some with minor differences. The main response was this: have the room/furniture of prototype 1 and the subtitles of prototype 2. Some also preferred the visible hands, prototype 2's buttons, or the removal of the backlog, but most seemed to think the room and the subtitles were the most important part of the experience.

Most of the open/general questions were answered without the participation knowing what the (features) the researcher was interested in researching. This means many of the features implemented were not or only very briefly touched upon by some participants. The remaining questions were thus designed to give the researcher a direct comparison between the two prototypes and the questions were a lot more specific to get clear answers from the participants. *What did you like more (and why)...*

...the buttons placed on the table of hovering in the air?

The responses to this question were the exact same: 'no preference'. Five participants said they had no preference at all, two said they slightly preferred the buttons to be on the table, as it would be a more natural place for them and one did prefer the buttons on the table, as it made more sense to the participant.

...the selected button always visible or only when an answer is available?

This question refers to the feature in prototype 1, where the selected button will remain visible until the new options arrive. From the eight participants, only four noticed this was happening and from them, one liked the feature, but thought the button needed to disappear slightly sooner and two thought the button should not stay, as it only generates irritation and confusion and is not that useful.

From the four users that did not notice it was a feature (which included the two target group participants, which is most likely due to the video not properly showing this feature), some useful data is gathered nonetheless. Three of them said it is probably better that the button disappears when you select it to avoid confusion.

...previous questions and answers visible or not?

To this question the responses were divided a bit more. Most proxy users felt it was not that useful, but it wasn't bothering them either as it was out of their peripheral vision. The others said it could be useful, especially for older adults that many have a harder time remembering. The target group participants both mentioned it to be useful if you forgot, or you just simply want to read it back. One interesting remark is that a feature where you can read the backlog after your session could be interesting and useful to some people, as a form of self-reflection.

...subtitles visible or not?

This question had the most simple responses. All participants unanimously agreed that subtitles were an important part of the system, for the obvious reasons of the system being more accessible, understandable and easier to use.

…’grab’ the buttons with the trigger or press the buttons with your index finger?

Overall, the participants’ preference leaned towards the pressing of the buttons instead of grabbing them, as this felt more intuitive and realistic (you also press a physical button). The button itself also gave more feedback and a greater sense of realism than the button in prototype 1. The only concern of the participants is that getting used to the controls for controlling your hands to press might require a lot of practice and might be difficult to get used to for older adults.

The target groups participants however, did not see or notice the difference in selecting a button. This is most likely due to the video going over it very quickly and not providing close enough attention and explanation on what is happening.

...the room with the table or the room without the table?

This question again provided very clear results, all participants preferred the room of prototype 1 (with the table). The cosiness it creates by its warm colours is very relaxing and makes the person feel at home. Some small things that could possibly improve the room is by removing the desk to make it more personal or by letting the coaches (and possibly the user) sit on a couch to make it less formal and more relaxed, according to the participants.

...the controllers in VR visible as controllers or visible as hands?

All proxy users indicated that they preferred the controllers to be visible as hands rather than controllers, as it is more interesting and realistic, which is good for engagement. One side note is that, as mentioned before, older adults might have difficulty getting used to the controls, which may require some training or require some changes to the current system to account for this.

The target group participants again did not see what the hands looked like in VR, just as the interaction, which is caused by the way the interactive video has been made and thus have no comments to add to this.

Besides the questions some additional remarks were made that did not fit as an answer to any of the questions. One participant argued that VR might not be the solution for COUCH and that a screen based version, where the user sits in front of three monitors and presses physical buttons, might work better. Another possible feature suggested by a participant is to add a ‘back’ button for when someone makes a mistake, or wants to change their answer. One other suggestion was that the system actively asks and tracks how someone's day was, to make sure their psychological state is also in good shape. Lastly, a couple participants wanted to let the researcher know that they really liked the experiment.

After this analysis was done, the results were integrated into the table proposed in the ideation phase (figure 5). For each feature the preferences of the participants, as described in this section, is colour coded. Light green means there is no clear preference and darker green means there is a clear preference. This was done to create an overview of the overall consensus, which puts everything in perspective and allows it to be used effectively in the conclusion chapter. Note: Darker green does not mean the feature of that prototype is better, it means that the clear majority of participants preferred it.

	Prototype 1	Prototype 2
Button Location Buttons are displayed in 3D and will be placed in a row next to each other.	**On a table** in front of the user.	Hovering **in the air** where they do not obstruct the view of the coaches.
Button Timing Whether the buttons are visible at all times or only when the user is prompted to answer.	**Always visible**, even when no questions are being asked.	**Only when prompted** and disappear when the question is answered.
Previous Question & Answers Whether the previously asked questions and given answers are being displayed somewhere on screen.	**Yes**, the previous questions & answers are displayed.	**No**, the previous questions and answers are not displayed.
Subtitles The display mode of the subtitles of the conversation.	**No** subtitles, can use previous questions and answers.	Placed in **worldspace** where it is not obstructive.
Interaction Mode Controllers are used to interact with the buttons.	**Grab** the buttons to select them.	**Push** the buttons to select them.
Location The location of the scene in which the conversation takes place.	A **cosy work room** that is warm in colour makes the users more at ease.	An **empty/clean room** that minimizes distractions.
Controllers appearance The controllers can appear as many different shapes in VR.	Appear **as controllers**, the same ones in VR as in real life.	Appear **as hands** instead of controllers.

Figure 23: Overview of features with preferences marked

7.3.3 Observations

Observations were made and noted during the proxy user participants' time with the prototypes. Most of the observations were later brought up by the participants in the interviews. The most notable observations were that the users needed a bit more time to get used to the controls in prototype 2, but within 10-20 seconds fully understood how it worked. Besides that the interaction in both versions went very smooth for all participants and no one needed extra instructions on how to use something. The only other important observation was that the buttons sometimes didn't work, but this was only the case for the first participant.

Chapter 8: Conclusion

This project aimed to find out if the Council of Coaches could get use out of implementing VR into their system and which features work best for this in terms of user engagement. Based on the findings in the literature in chapter 2 and the results summarised in section 7.4 of the evaluation chapter, the sub-research questions and the main research question will be answered:

"What are the benefits/downsides of using VR in applications with ECAs?"

Improved user engagement is the largest benefit for any system that requires the users' attention. An improved user engagement brings many benefits along with it, such as having a better mood and more credibility. The biggest downside is the cost of production and supplying every person with headsets, besides that VR can cause problems for some people, such as motion sickness.

"What are the VR applications/features that work best for older adults?"

Having limited movement and short sessions would most likely work best. This should be combined with a simplistic and easy to understand interaction system, so that the barrier of entry older adults can experience with new technologies, is lowered. The interaction should feel intuitive and logical, while the controls in VR are often not a problem due to them being so intuitive. From the user tests it can also be concluded that the environment should be somewhere where the user feels comfortable and subtitles are an absolute necessity to provide clarity and accessibility.

"What are the best ways to interact with a VR system using conversational agents for older adults?"

The results from the target group interviews cannot be used to answer this question, as the interactive video did not show these features well enough. For the proxy users however, it can be concluded that clicking a button in VR that resembles clicking a button in real life would be a good solution for the target group.

From all literature and results gathered and the conclusions from the sub-research questions the main research question can be answered.

“What is the effect on user engagement when implementing VR in the Council of Coaches system?”

Between the two created prototypes the engagement was measured for various categories. Scores on engagement were high, with proxy users giving overall higher scores than online users. From research findings it was expected the VR versions would score higher in engagement and this is especially evident when looking at the interaction, where the interaction of the proxy users was far more elaborate than the online users, resulting in higher engagement scores in PU, as can be seen in figure 23. Neither of the two versions can be concluded as better, but prototype 1 does score slightly higher in terms of engagement than prototype 2. Its appearance had the most impact on its score and AE is also the area that showed the clearest difference between the two prototypes. This, and other conclusions, show that the appearance (cosy room) and accessibility (subtitles) are the most important factors for user engagement in the Council of Coaches system. Thus, the ideal version, as agreed upon by all participants, is having a cosy, warm room and subtitles to support the user.

Not all features tested were noticeably impactful to the user engagement, but the preferences of the users can still be utilised and from this some basic ‘guidelines’ can be established:

- Button location: No clear results
- Button timing: Only show buttons when prompting for answers
- Previous Q&A: Useful for older adults, keep to the side to avoid distraction
- Subtitles: Necessity, should be clearly readable and unobstructive
- Interaction mode: Push buttons, add enough feedback
- Location: Cosy room, should be warm and comforting
- Controllers: Hands, should use simple controls

Chapter 9: Discussion

This chapter will reflect on the project process and discuss the various decisions made and look at possibilities for future work.

Online target group interviews & the interactive video

The first interesting point of discussion are the target group interviews. In a perfect scenario the physical VR tests would have taken place using older adults, however as this was not possible an interactive video was used instead. While this had many advantages over a normal video, VR is still a very difficult thing to display using another medium, especially for those who have never experienced VR. As feared, the result of this is that the online interviews did not yield good data in some areas, such as interaction. The results of the questionnaire are somewhat similar however, but the sample size was too small to exclude the possibility of it being chance. The interactive video also took very long to make and it is interesting to see whether a normal video would have yielded the same results.

No user confrontation

Often in research during the ideation phase, a user confrontation takes place. This happens to assess whether the ideation is going in the right direction and to get useful feedback. Due to the pandemic and the abstraction of this subject, this confrontation did not take place. This is something that could have influenced the design phase and could be explored in later iterations of the project.

Realisation problems

The realisation of the prototypes was plagued by problems. The research was completely unfamiliar with the software in use and COUCH can be very overwhelming. A lot of time was lost trying to understand existing scripts and figuring out how to edit the dialog. With the amount of time lost in mind, it might have been more time efficient to build a new system from scratch that mimics the basic functionality of COUCH, as the real things that make COUCH special, multiple agent interaction and argumentation frameworks, were used in the prototypes.

Evaluation

The evaluation as performed in this research might not be the optimal way. Other techniques or questionnaires could be used to gain different results. The participants were also (due to the pandemic) all recruited by using convenience sampling. As the participants personally knew the researcher this could have an impact on the credibility of the results.

Future work

While this research focussed on what features work best in VR for COUCH, no baseline was established with a screen based version. For future research, a proper screen based version could be designed and then tested versus the VR version, to quantify the added benefit of VR, for COUCH specifically.

Also, (only) seven features were developed and tested for this project and many more can be designed for this project, maybe even some big changes to the interaction system (not using buttons).

As mentioned before, the target group could not test with the VR installation, this is something that should be explored in the future, as proxy users are never a perfect substitute. As the proxy users also knew the researcher, the results could have been skewed and thus research with random participants could provide valuable information.

References:

[1] World Health Organisation (WHO). *Population ageing.* https://www.who.int/features/qa/72/en/ 2010

[2] H. op den Akker, R. op den Akker, T. Beinema, O. Banos, D. Heylen, B. Bedsted, A. Pease, C. Pelachaud, V. Traver-Salcedo, S. Kyriazakos and H. Hermens, "Council of Coaches – A Novel Holistic Behavior Change Coaching Approach", in Proceedings of the 4th International Conference on Information and Communication Technologies for Ageing Well and e-Health – Volume 1: ICT4AWE (pages 219-226), Funchal Madeira, Portugal, March 2018.

[3] Council of Coaches. https://council-of-coaches.eu/ 2020

[4] J. Lester, K. Branting, B. Mott. "Conversational agents". The practical Handbook of internet computing. 2004

[5] R. B. Kantharaju, D. De Franco, A. Pease, and C. Pelachaud. "Is Two Better than One?, Effects of Multiple Agents on User Persuasion." In IVA '18: International Conference on Intelligent Virtual Agents (IVA '18), (Sydney, NSW, Australia. ACM, New York, NY, USA). 2018

[6] L. Coa, C. Peng, J.T. Hansberger. "Usability and engagement study for a serious virtual reality game of lunar exploration missions". 2019

[7] D. Allcoat, A. von Mühlenen. "Learning in virtual reality: Effects on performance, emotion and engagement". 2018

[8] C.X. Lin, C. Lee, D. Lally & J. F. Coughlin. "Impact of virtual reality (VR) experience on older adults' well being". 2018

[9] J.F. Morie, E. Chance, K. Haynes, D. Rajpurohit. "Embodied conversational agent avatars in virtual worlds: Making today's immersive environments more responsive to participants" in *Can Computers Play Like People?*, vol 9783642323232. Playa Vista, California, USA: Springer-Verlag Berlin Heidelberg, 2012, pp. 99-118.

[10] T. Pesja, M. Gleicher, B. Mutlu. "Who me? How virtual agents can shape conversational footing in virtual reality". 2017

[11] Visualise. *Virtual Reality in Healthcare.* https://visualise.com/virtual-reality/virtual-reality-healthcare

[12] A. Lele. "Virtual reality and its military utility". J Ambient Intell Human Comput 4, 17–26. 2013

[13] D. Vajak and Č. Livada. "Combining photogrammetry, 3D modeling and real time information gathering for highly immersive VR experience." Paper presented at the 2017 Zooming Innovation in Consumer Electronics International Conference: Galvanize Your Creativity, ZINC 2017, 82-85. 2017

[14] I. Mahalil, M. E. Rusli, A. M. Yusof, M. Z. M. Yusoff and A. R. R. Zainudin, "Study of immersion effectiveness in VR-based stress therapy," Proceedings of the 6th International Conference on Information Technology and Multimedia, Putrajaya, 2014, pp. 380-384. 2014

[15] M.K. Holden. “Virtual environments for motor rehabilitation: Review”. Cyberpsychology and Behavior, 8(3), 187-211. 2005

[16] G. Ye and C. Liu, "Application of VR in appendectomy surgery system," 2010 3rd International Conference on Biomedical Engineering and Informatics, Yantai, 2010, pp. 118-121. 2010

[17] D. Freeman, S. Reeve, A. Robinson, A. Ehlers, D. Clark, B. Spanlang, & M. Slater. “Virtual reality in the assessment, understanding, and treatment of mental health disorders”. Psychological Medicine, 47(14), 2393-2400. 2017

[18] J. Varela-Aldás, E. M. Fuentes, G. Palacios-Navarro, and I. GarcíaMagariño, “A comparison of heart rate in normal physical activity vs. immersive virtual reality exergames,” in Human systems engineering and design II (T. Ahram, W. Karwowski, S. Pickl, and R. Taiar, eds.), vol. 1026 of Advances in intelligent systems and computing, pp. 684–689, Cham: Springer International Publishing, 2020.

[19] Nuance. *Nina.* https://nuance.com/omni-channel-customer-engagement/digital/virtual-assistant/nina

[20] W. Swartout et al. “Ada and Grace: Toward Realistic and Engaging Virtual Museum Guides” 2010

[21] Aetna. *vHeatlh.* https://vhealth.io

[22] S. Pedram, P. Perez, S.A. Palmisano & M. Farrelly. “A Systematic Approach to Evaluate the Role of Virtual Reality as a Safety Training Tool in the Context of the Mining Industry”. 2016

[23] M. Taillade, H. Sauzéon, M. Dejos, P. Arvind Pala, F. Larrue, G. Wallet, C. Gross, B. N'Kaoua. “Executive and memory correlates of age-related differences in wayfinding performances using a virtual reality application”. Aging, Neuropsychology, and Cognition, 20(3), 298-319. 2013

[24] K.J. Miller, B.S. Adair, A.J. Pearce, C.M. Said, E. Ozanne, M.M. Morris. “Effectiveness and feasibility of virtual reality and gaming system use at home by older adults for enabling physical activity to improve health-related domains: a systematic review”. Age and Ageing, Volume 43, Issue 2, Pages 188–195. 2014

[25] N. Skjæret, A. Nawaz, T. Morat, D. Schoene, J.L. Helbostad, B. Vereijken. “Exercise and rehabilitation delivered through exergames in older adults: An integrative review of technologies, safety and efficacy”. International Journal of Medical Informatics. Volume 85, Issue 1, Pages 1-16. 2016

[26] A. Mirelman, L. Rochester, I. Maidan, S. Del Din, L. Alcock, F. Nieuwhof, M.O. Rikkert, B.R. Bloem, E. Pelosin, L. Avanzino, G. Abbruzzese, K. Dockx, E. Bekkers, N. Giladi, A. Nieuwboer, J.M. Hausdorff. “Addition of a non-immersive virtual reality component to treadmill training to reduce fall risk

in older adults (V-TIME): a randomised controlled trial". The Lancet. Volume 388, Issue 10050, Pages 1170-1182. 2016

[27] World Health Organisation (WHO). *eHealth at WHO*. https://www.who.int/ehealth/about/en/

[28] P. Lindner, A. Miloff, E. Zetterlund, L. Reuterskiöld, G. Andersson, & P. Carlbring. "Attitudes toward and familiarity with virtual reality therapy among practicing cognitive behavior therapists: A cross-sectional survey study in the era of consumer VR platforms". Frontiers in Psychology. 2019

[29] G. Riva, M. Bacchetta, G. Cesa, S. Conti, & E. Molinari. "E-health in eating disorders: Virtual reality and telemedicine in assessment and treatment". 2002

[30] W.A. Turner, L.M. Casey. "Outcomes associated with virtual reality in psychological interventions: where are we now?". Clinical Psychology Review, Volume 34, Issue 8, Pages 634-644. 2014

[31] A.H. Mader & W. Eggink. A Design Process for Creative Technology. In *Proceedings of the 16th International conference on Engineering and Product Design, E&PDE 2014* (E. Bohemia, A. Eger, W. Eggink, A. Kovacevic, B. Parkinson, & W. Wits, eds.) pp. 568-573. The Design Society. 2014

[32] Valve. *Half-life alyx.* https://www.half-life.com/en/alyx/

[33] V. Mitousia and O. Giotakos. "Virtual Reality therapy in anxiety disorders". 2016

[34] U.S. Department of Health & Human Services. *User Interface Design Basics.* https://www.usability.gov/what-and-why/user-interface-design.html

[35] Oculus. *Introduction to best practices.* https://developer.oculus.com/design/bp-vision/

[36] UX Planet. *Accessible Design: Designing for the Elderly.* https://uxplanet.org/accessible-design-designing-for-the-elderly-41704a375b5d

[37] J. Jerald. "Input Devices" in *The VR Book: Human-Centered Design for Virtual Reality*. Association for Computing Machinery and Morgan & Claypool. 2015, ch.27, pp. 307-322

[38] S. Fröjdman. "User Experience Guidelines for Design of Virtual Reality Graphical User Interfaces". 2016

[39] H. O'Brien, P. Cairns & M. Hall. "A Practical Approach to Measuring User Engagement with the Refined User Engagement Scale (UES) and New UES Short Form." International Journal of Human-Computer Studies. 112. 10.1016/j.ijhcs.2018.01.004. 2018

A Miscellaneous

A1 - Ideas

Vr implementation

- UI elements
- Interaction
- Visualization

UI elements

- No buttons/options
  - Voice recognition
  - Keyboard
- Dropdown menu
- Selection
  - **Light up/change colour**
  - Size increase
  - No difference
- Full screen overlay
  - Flat UI
  - Curved UI (towards user)
  - Large buttons
  - Divide screen in options
- World space buttons
  - Position
    - In front of coaches (far from user)
    - Under coaches (far from user)
    - In the air (close to player)
    - **On table (close to player)**
    - Stacked vertically (or above one another)
    - **Stacked horizontally (or next to one another)**
  - Orientation
    - Forward
    - **Towards player**
    - **Upwards (when on table perhaps)**
  - Size
    - Very big (far away)
    - **Readable (close)**
    - Small
  - Order
    - Random (or as the system provides them)

            - **Always same order (positive message always same place)**
    - Colour
        - All the same colour
        - **Same kind of message same colour (positive = blue)**
- Overlay buttons
    - Stacked vertically
    - Stacked horizontally
    - Flat UI (buttons facing forward)
    - Curved UI (buttons facing user)
- Buttons
    - Always visible -> remember last question
    - **Only when prompted**
- Subtitles
    - World space
        - Floating above speaker
        - Below speaker
        - Centralized below/in front
            - No difference between speakers
            - Colour difference when speaking (e.g. carl = red)
            - Attach names to subtitles (e.g. “Carl: SENTENCE”)
    - **Overlay**
        - **At bottom of screen**
        - At top of screen
        - Appearance
            - No difference between speakers
            - Colour difference when speaking (e.g. carl = red)
            - **Attach names to subtitles (e.g. “Carl: SENTENCE”)**
            - Flat
            - **Curved**

Interaction: User -> System

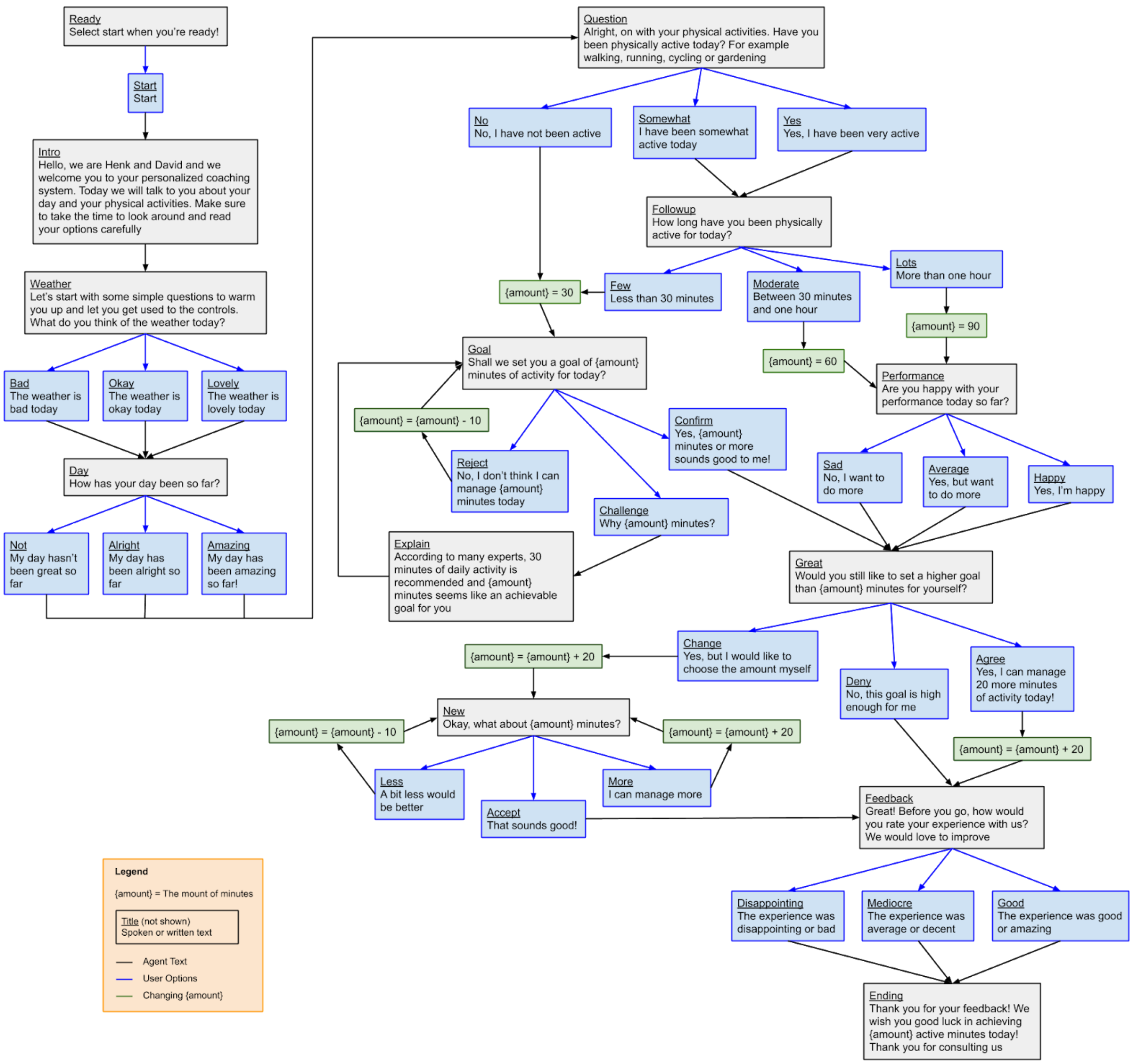
- **Grab options with controllers**
    - Floating list
    - **Floating buttons**
- Swipe with controller
    - Swipe between options (put correct one in middle)
    - Swipe correct option up
- Options attached to a controller
    - Use buttons to select or scroll through list/menu
    - Using the other controller to select
- Map options to buttons on controllers
    - Use ‘A’ to select the first option, ‘B’ for second
    - User joystick to cycle between options

- Aim at options to trigger (time based)
    - Controller (laser appears to show where user is pointing)
        - Aim at option for long enough
        - Aim at option then press trigger to confirm
    - Headset
        - ‘Stare’ at option long enough to trigger
        - Add dot in screen to see the middle
        - No controllers needed
- Voice recognition
    - Headset microphone
    - 3rd party microphone
- Use real hands (using e.g. leap motion)
    - Hands as controllers

Visualization

- Location
    - Conference room
    - Doctors office
    - **Normal office (with bookshelves)**
    - Outside?
    - Empty/clean room
- Agents
    - **Default appearance (difference in colour clothing and different face & voice)**
    - Tailored clothing to their profession/expertise
- User
    - **No visual representation**
    - Static visual representation (only legs and torso)
    - Dynamic visual representation (IK)
- Controllers
    - Appear as controllers
    - **Appear as hands**
    - Appear as other object
        - Laser pointer
        - Ball
        - Arrow/cursor
        - Don’t appear

A2 - Dialog tree



B User tests

B1 - Information brochure

This research will revolve around the implementation of virtual reality in the Council of Coaches system.

COUNCIL of COACHES

Background

The Council of Coaches (or COUCH) is an interactive coaching system that uses multiple embodied conversational agents that can both interact with the user, as well as with each other. The team behind COUCH try to push the state of the art in conversation agents and provide older adults with an integrated coaching solution.

COUCH provides coaching by having multiple coaches, each with their own expertise, giving health and wellbeing related advice to the users, increasing the availability and quality of health care.

This research aims to see whether Virtual Reality can be a useful tool for COUCH and what features (in VR) would work the best for the system. Multiple versions with different types of interactions and features will be created and tested in order to accomplish this.

Research procedure

Interviews, questionnaires and user tests will be conducted in June 2020 to gather data needed to conclude the research.

User tests will be conducted in which the participant will be presented with a couple different versions of the interaction system of COUCH in virtual reality and asked to use each version for a few minutes and complete a dialog with the coaches by choosing the options they want. After this their opinion will be recorded using a semi-structured interview. In this interview the user will be asked which version and features they liked/disliked and why, and also what their 'perfect' version of the system would look like. Video recordings of the user test can be used to later analyze how the users reacted to different features. The user test and the interview will take approximately 30 minutes total.

Due to COVID-19 the target group of older adults is not available for face-to-face user testing. In order to get input from the target group, some online interviews will be conducted. These interviews will show the participant the different versions as pre recorded videos of interaction with the system (as the participants do not have the equipment needed at home) and will then ask for their opinion on the different solution in a semi-structured interview. These interviews will be recorded to be used for later analysis and will take about 20 minutes.

Personal Data

All the data that is gathered will solely be used for research purposes, will never be shared with third parties and will be anonymous if made public.

The personal data will be stored securely and according to the VSNU guidelines will be stored at least 10 years. It will only be accessed by the researchers working on the project.

Participation

The participation in this research is completely voluntary and the participant can stop at any point in time, without needing a reason and without consequences.

The participant, for the face-to-face user test, should be in good health and be aware of the possible side-effects virtual reality can cause for some people, e.g. motion sickness, epileptic seizures, etc.

Contact

If during or after the research any question or need for additional information arises, the participant can contact the researchers by emailing Rens van der Werff (r.vanderwerff@student.utwente.nl) or his supervisor Randy Klaassen (r.klaassen@utwente.nl)

If the participant would like to receive independent advice or file a complaint, the ethics committee of EEMCS can be contacted (ethics-comm-ewi@utwente.nl).

UNIVERSITY OF TWENTE.

B2 - Informed consent form

Informed consent form

About

This research aims to see whether Virtual Reality can be a useful tool for the Council of Coaches (COUCH) system and what features (in VR) would work the best, as explained in the provided information brochure "Implementing Virtual Reality in the Council of Coaches system".

Contact

If during or after the research any question or need for additional information arises, the participant can contact the researchers by emailing Rens van der Werff (r.vanderwerff@student.utwente.nl) or his supervisor Randy Klaassen (r.klaassen@utwente.nl)

If the participant would like to receive independent advice or file a complaint, the ethics committee of EEMCS can be contacted (ethics-comm-ewi@utwente.nl).

Research

I hereby declare the following by signing this form:

- ❑ I am fully informed about the research and the purpose and method of the research is clear to me. I have had the opportunity to ask questions and my questions have been answered sufficiently.
- ❑ I am aware that my participation in the research is voluntary and that I have the right to withdraw my consent at any moment, without consequences.
- ❑ I consent to the gathering of anonymous information and video recordings during the research. The video recordings will only be used by the researchers to gather more data and will never be made public or shared with third parties. Other data/results will always be used anonymously.
- ❑ I allow my answers to questions to be anonymously quoted.
- ❑ I declare that I am in good health and am aware that Virtual Reality could induce symptoms such as motion sickness and epileptic seizure.

Date . . . .

Name . . . .

Signature . . . .

B3 - Introduction

A short introduction told at the beginning of a user test.

Proxy users

Welcome and thank you for participating in the user test of my bachelor thesis. For my thesis I have created two prototypes that I would like your opinion on to see if and how Virtual Reality can be applied in the Council of Coaches system. This system aims to provide personalized health and wellbeing coaching to older adults by means of an interactive coaching application. As this system is being designed for older adults, for this test I would like you to answer all questions as if you were an older adult (60+). You can, for example, imagine how your (grand)parents would react to such a system.

Before using each prototype, you will receive a short instruction on how to use the given prototype and after you have used one, you will get a short questionnaire. After the second questionnaire there will be an interview to discuss your opinions on the two prototypes. Once finished you will receive a candybar. To start I need you to read and sign the consent form.

Online target group users

Welcome and thank you for participating in the user test of my bachelor thesis. For my thesis I have created two prototypes that I would like your opinion on to see if and how Virtual Reality can be applied in the Council of Coaches system. This system aims to provide personalized health and wellbeing coaching to older adults by means of an interactive coaching application.

You will be shown an interactive video of both prototypes, but before each prototype you will receive a short instruction on how to use the given prototype and after you have used one, you will get a short questionnaire. After the second questionnaire there will be an interview to discuss your opinions on the two prototypes. To start I need you to read and sign the consent form.

B4 - Instructions

Before each prototype is shown/used, an instruction will be given to help users on their way. The first part of the instruction is a general instruction which will be given to all participants, but only before the first prototype.

General instruction

In a moment you will see/be seated in a room where there are two virtual coaches. These coaches will have a conversation with you and ask you various questions, this conversation will be the same for both prototypes. You answer these questions by selecting an answer from the options on the buttons that appear. Take your time to look around and read the questions and answers. The conversation will notify you when you are done.

Proxy users - Concept 1

In this prototype you need to select the buttons by moving your controller into one and pressing the trigger at your index finger.

Proxy users - Concept 2

In this prototype you can select a button by 'pressing' it with your index finger. Your hand in VR will respond to your grip on the controller, meaning you can point with your index finger by squeezing all the other fingers.

Online target group users

After each question you can select the option of your choice by clicking on it with the mouse.

B5 - Short Questionnaire

The questionnaire was made in Google Forms without using any participant data, instead, the participants are numbered so that the responses to both prototypes can be compared.

Questions

The first two questions are filled in by the researcher:

- *Participant number*
- *Prototype number*

The next 12 questions all use the following scale as their answers:
strongly disagree - disagree - neither agree nor disagree - agree - strongly agree.
These question (as proposed by 0'Brien et al. [39]) are:

- *I lost myself in this experience*
- *The time spend using this prototype just slipped away*
- *I was absorbed in this experience*
- *I felt frustrated while using this prototype*
- *I found this prototype confusing to use*
- *Using this prototype was taxing*
- *This prototype was attractive*
- *This prototype was aesthetically appealing*
- *This prototype appealed to my senses*
- *Using this prototype was worthwhile*
- *My experience was rewarding*
- *I felt interested in this experience*

B6 - Interview

The interview questions will focus on the features and their realisation in both prototypes, as well as the prototypes as a whole.

List of features:

Button Location Buttons are displayed in 3D and will be placed in a row next to each other.
Button Timing Whether the buttons are visible at all times or only when the user is prompted to answer.
Previous Question & Answers Whether the previously asked questions and given answers are being displayed somewhere on screen.
Subtitles The display mode of the subtitles of the conversation.
Interaction Mode Controllers are used to interact with the buttons.
Location The location of the scene in which the conversation takes place.
Controllers appearance The controllers can appear as many different shapes in VR.

The interview questions can be found on the next page.

General questions:

- *What did you think of the coaches?*
- *What did you think of the dialog?*
- *What did you think of prototype 1 overall?*
- *What did you think of prototype 2 overall?*
- *Which prototype did you like the most overall and why?*
- *Which prototype do you think would work best and why?*
- *What would be the ideal version (mixing features)?*

Questions about all the features: *What did you like more (and why)...*

- *...the buttons placed on the table of hovering in the air?*
- *...the selected button always visible or only when an answer is available?*
- *...previous questions and answers visible or not?*
- *...subtitles visible or not?*
- *…'grab' the buttons with the trigger or press the buttons with your index finger?*
- *...the room with the table or the room without the table?*
- *...the controllers in VR visible as controllers or visible as hands?*

C Results

C1 - Questionnaire results

The questionnaire results without the pilot tests, coloured background to distinguish prototype versions, coloured text to distinguish participants. The table is ordered by participant and the average for each combination of participant and version is provided in the last column.

n	Type	FA-1	FA-2	FA-3	PU-1	PU-2	PU-3	AE-1	AE-2	AE-3	RW-1	RW-2	RW-3	**Avg**
Proxy user test														
1	1	4	4	4	5	5	2	4	4	5	3	2	5	**3.9**
1	2	3	3	3	5	5	5	2	2	3	4	4	1	**3.3**
2	2	4	3	4	5	5	5	4	4	4	4	3	5	**4.2**
2	1	4	4	4	5	4	5	5	5	4	4	4	5	**4.4**
3	1	3	4	4	5	4	3	4	4	4	3	4	5	**3.9**
3	2	5	4	4	5	4	4	4	4	5	4	4	5	**4.3**
4	2	3	3	2	2	4	4	3	4	3	3	2	3	**3.0**
4	1	2	2	4	4	4	4	4	4	3	3	3	3	**3.3**
5	1	3	4	4	4	4	4	3	3	4	4	3	4	**3.7**
5	2	4	4	4	4	3	4	4	3	5	4	4	5	**4.0**
6	2	2	2	2	4	3	4	2	2	3	2	2	2	**2.5**
6	1	4	3	4	4	4	3	5	5	4	4	3	4	**3.9**
Online target group test														
7	1	4	4	4	3	3	4	5	5	4	4	4	4	**4.0**
7	2	4	4	4	4	4	4	3	3	2	3	4	4	**3.6**
8	2	2	3	2	3	3	4	4	3	3	3	3	4	**3.1**
8	1	3	3	2	4	4	4	4	4	4	4	3	4	**3.6**