

THE POWER OF A HUMAN-CENTERED VR EXPERIENCE OF THE FUTURE OF WORK

Master Thesis by Pak Long Cheung

September 2021

TU Delft, Faculty of Industrial Design Engineering
M.Sc. Integrated Product Design

Colophon

This is the final report of my graduation project
'The power of a human-centered VR experience of the
future of work'.
Executed as a research design project for the Faculty
of Industrial Design Engineering in Delft, this project
completes my Master's program Integrated Product
Design.

Author: Pak Long Cheung

September 2021

Master Thesis
M.Sc. Integrated Product Design
Faculty of Industrial Design Engineering, Delft
University of Technology

Supervisory Team:

Chair: Dr. Aschenbrenner, D.
Mentor: Henny, J.

Delft University of Technology
Faculty of Industrial Design Engineering
Landbergstraat 15
2628 CE Delft
The Netherlands

LIST OF ABBREVIATIONS:

AGV	Automated Guided Vehicles
AR	Augmented Reality
FoW	Future of Work
GUI	Graphical User Interface
HMD	Head Mounted Device
JTD	Job to be done
MR	Mixed Reality
PBL	Points, Badges and Leaderboards
PQ	Presence Questionnaire
SSQ	Simulator Sickness Questionnaire
SUS	System Usability Scale
TUD	University of Technology, Delft
UX	User Experience
VR	Virtual Reality

Table of contents

INITIATE	INTRODUCTION	6	FUTURES
EXPLORE	CONTEXTUAL STUDY	8	ROBOHOUSE
DISCOVER	LITERATURE STUDY	10	VR EXPERIENCE
DEFINE	LITERATURE STUDY	14	PRELIMINARY ANALYSIS
(RE)DEVELOP	LITERATURE STUDY	20	METHOD
DEVELOP	LITERATURE STUDY	22	PRESENCE & INTERACTIVITY
(RE)DEVELOP	SYNTHESIS	30	IMMERSION
DEVELOP	SYNTHESIS	36	CONCLUSION
(RE)DEVELOP	SYNTHESIS	39	LIMITATION
DEVELOP	SYNTHESIS	40	IMMERSION CRITERIA & ELEMENTS
(RE)DEVELOP	SYNTHESIS	44	CURRENT & DESIRED SCENARIO
DEVELOP	SYNTHESIS	46	HUMAN-CENTERED APPROACH
(RE)DEVELOP	SYNTHESIS	48	BRAINSTORM
DEVELOP	SYNTHESIS	49	THE RITE METHOD
(RE)DEVELOP	SYNTHESIS	56	ITERATIONS
DEVELOP	SYNTHESIS	76	RECAP
VERDICT	CONCLUSION	82	VR FOR FORESIGHT
VERDICT	CONCLUSION	83	PRESENCE STUDY
VERDICT	CONCLUSION	88	VR FOR RESEARCH

Introduction

FUTURES

Futures is a topic that is inherently fuzzy, as everything beyond this moment can be a potential future. The uncertainty that is related to futures makes it not only difficult to imagine, but also challenging to choose appropriate forward-looking disciplines. Dorsser et al. (2018), bridged this challenge by connecting the types of futures identified by Voros (2003) and the various related levels of uncertainties (defined by Walker et al., 2013). Because forward-looking disciplines often target a specific type of future, the resulting framework can be used as a “practical guide for futures researchers on what methods to use when anticipating future developments at specific levels of perceived uncertainty” (Voros, 2003, 2017).

FUTURE TYPES

Figure 1 is the “Futures Cone” developed by Voros (2003, 2017). The cone illustrates the different types of futures, based on where we are now and the likelihood of future developments (i.e., certainty/uncertainty). The cone also illustrates how the different types of futures relate to each other.

As earlier mentioned, the potential future encompasses everything beyond the current moment. Four main types of potential futures were identified by Voros and used by Dorsser et al.: potential possible future, potential plausible future, potential probable future and the potential projected future. The cone also illustrates the potential preferable future, which is a subset of the potential futures that is ‘desirable’. This type of future was omitted by Dorsser et al. during construction of the guide, as classifying issues according to Walker’s structure does not take into account futures in terms of preferences, but according to uncertainties.

The four main types of futures are described as follows by Dorsser et al.:

- Potential *Possible* Future — “those futures that we think ‘can’ happen, based on some future knowledge we do not yet possess, but which we might possess someday (e.g., warp drive)” (Dorsser et al, 2018).
- Potential *Plausible* Future — “those futures that we think ‘could’ happen based on our current understanding of how the world works (physical laws, social processes, etc.)” (Dorsser et al, 2018).
- Potential *Probable* Future — “those futures that we think are ‘likely to’ happen, based on current trends” (Dorsser et al, 2018).
- Potential *Projected* Future — “the (singular) default, business as usual, ‘baseline’, extrapolated ‘continuation of the past through the present’ future, that could also be considered as being ‘the most probable’ (or most expected) of the Probable futures” (Dorsser et al, 2018).

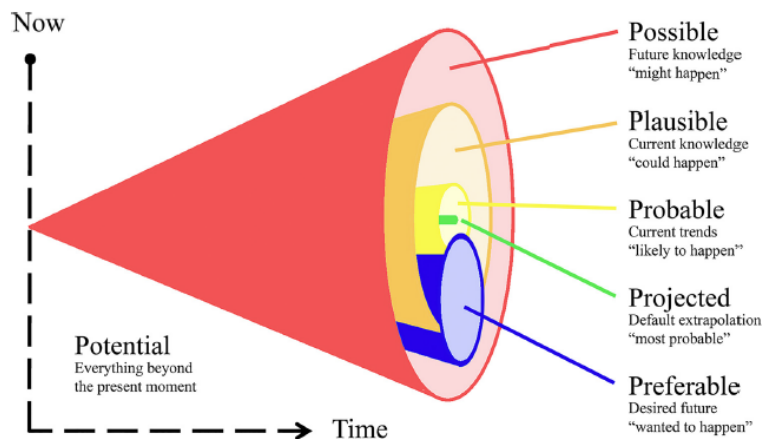


FIGURE 1. FUTURES CONE (VOROS, 2003)

Dorsser et al. concluded that the core focus of several forward-looking disciplines can be linked to the type of future (as defined by Voros) it tries to address. When further linking the forward-looking disciplines with the various level of uncertainties (as defined by Walker et al., 2013), a ‘Futures Pyramid’ (figure 2) is imagined where an ‘appropriate’ forward looking discipline is suggested, based on the level of uncertainty in the type of future addressed.

The pyramid consists out of four layers of anticipating future developments. The top of the pyramid represents the little knowledge we possess and is associated with the highest level of uncertainty. Similarly, the base of the pyramid represents a future that is relatively certain and is associated with the lowest level of uncertainty. The layers; types of futures and related uncertainties, together with the recommended approaches, are structured as follows:

- "Layer 1: Deterministic Forecasting links to projected futures and Level 1 uncertainty";
- "Layer 2: Probabilistic Forecasting links to probable futures and Level 2 uncertainty";
- "Layer 3: Foresight links to plausible futures and Level 3 uncertainty";
- "Layer 4: Futures links to possible futures and Level 4 uncertainty" (Dorsser et al., 2018).

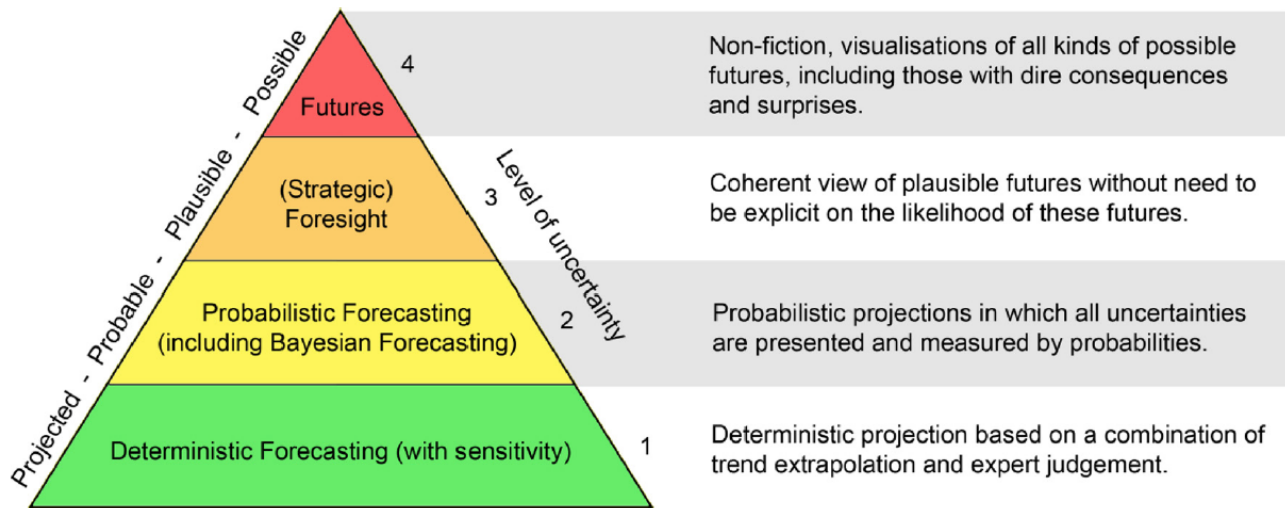


FIGURE 2. FUTURES PYRAMID (DORSSER ET AL., 2018)

FUTURE OF WORK

When talking about the future of work, 'Industry 4.0' is a concept that is mentioned often. Industry 4.0 is a wide concept of the future of work where physical operation and production is enhanced by making use of different smart, assistive and analytical technologies. Although much can be found on this topic currently, the debate on the specifics and also the time-frame of implementation (of industry 4.0) is still ongoing.

Uncertainty is affected by the time of implementation; the farther down in the future the implementation lies, the more uncertainties will arise. Therefore, uncertainties are part of the concept, as it is difficult to state when the transition to Industry 4.0 might happen. Furthermore, depending on the expert asked, specifics of the concept vary and different scenarios can be painted.

Following the Futures Pyramid guide, including Voros' definition of types of futures; Industry 4.0 is a concept that paints a potential plausible future. Consequently, the guide suggests for foresight as an appropriate forward-looking discipline in order to "paint a coherent view". Specifically; "to develop a coherent forward view of plausible futures without the need to be explicit

on the probabilities with which these futures occur" (Dorsser et a., 2018).

A scenario of industry 4.0, where the emphasis is put on the collaboration between human and robots in industries, rather than the complete automation of it for instance, has been developed using Virtual Reality (VR). In this VR experience, workers assemble wheels onto a bicycle frame with the help of augmented reality (AR), by collaborating with various robots. Labor-intensive aspects of assembly in this scenario will mostly be handled by the robots. AR is used to guide the worker and support the teamwork. Although the VR experience is not commercially available, it can be experienced in the EXPO room of Robohouse.

Early usability test results (conducted with TU Delft and Robohouse, prior to the start of this project) has indicated players have troubles getting through the VR experience. Using the experience as foresight effectively, requires the potential plausible future of work scenario to be conveyed effectively. As players are currently occupied with the controls due to usability issues, it is questionable to what degree the VR experience is able to immerse the players into the scenario.

Contextual study



FIGURE 3. ROBOHOUSE

ROBOHOUSE

Robovalley believes that the development of (cognitive) robotics will help contribute to solving major societal challenges like ageing population and climate change. In manufacturing specifically, they mention how the application of AI and intelligent machines can help the process of reshoring production facilities, creating jobs and economic growth. This also helps to put an end to dull and dangerous jobs. Robovalley thinks that the Netherlands is a country that is well suited for solving these challenges, as it constitutes of highly educated professionals who are used to working together; which is a requirement for creating interdisciplinary solutions.

In order to drive solving these challenges, three initiatives were established. Robohouse (figure 3) is one of these three and is dubbed their 'Smart Industry Fieldlab'. They are closely linked to the University of Technology in Delft (TUD); organisationally, but also physically, as they are situated on the campus of TUD, near the faculty of Architecture. Robovalley further consists out of a vibrant network of general robot enthusiasts and project partners, and a start-up community.

Robohouse has several different facilities. A big expo room (figure 4) where different projects (like the VR experience) and prototypes are situated, a test center which can provide as an industrial setting to conduct tests in, a 'learn' center, a make space and several different (project) rooms. Robohouse is a place where "companies can discover, develop & test advanced cognitive robotics applications" (Robovalley, 2020). A couple of recent projects Robohouse was involved in can be seen on the next page.



FIGURE 4. EXPO ROOM

HEINEKEN BOTTLE PICKER

60.000 bottles are created at bottle lines by Heineken. Some of these falls, and some even break. Depending on whether the bottle is pasteurized or not, the bottles or shards are hot. This creates dangerous situations for operators working at the bottling lines.

A robotic solution (figures 5 and 6) has been developed by Robohouse and its community members and Heineken, together with TNO, SMITZH (RoboValley, 2021a).

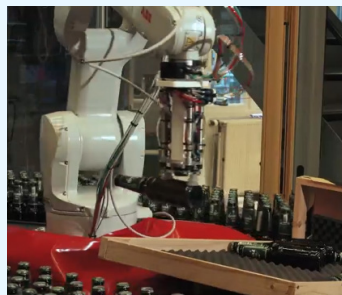


FIGURE 5. HEINEKEN ROBOT ARM

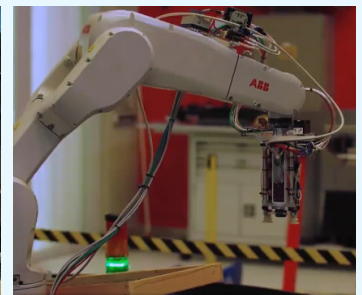


FIGURE 6. HEINEKEN ROBOT ARM

ROBOTICS MINOR

Together with several different partners, an educational program at the TU Delft has produced several solutions (figures 7 and 8) in order to “improve daily life”. In this robotics minor, students worked in interdisciplinary teams to research and provide for robotic solutions (RoboValley, 2021b).



FIGURE 7. BEACH CLEANER

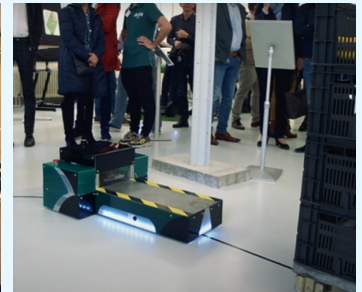


FIGURE 8. SUPERMARKET HELPER

PUTTING WHEELS ON A BICYCLE (VR EXPERIENCE)

Futures are uncertain and often hard to imagine. Together with TU Delft, Robohouse and its partners have sketched a scenario of how an assembly line in a future bicycle factory could look like (figures 9 and 10). As alluded to earlier, the experience can be played in the expo room of Robohouse (figure 4)

The VR experience is, like many of the other projects Robohouse is involved in, not commercially accessible. Due to its constricted context of use (the EXPO room in Robohouse; next to the faculty of Architecture on the TU Delft campus), relevant users can be relatively easily deduced.



FIGURE 9. VR EXPERIENCE



FIGURE 10. VR EXPERIENCE

Different projects surrounding robotic solutions and futures have been developed and tested by Robohouse together with various other individuals and organizations. These include students and community members, universities, research organizations and small-medium enterprises. Because of the (restricted) context of use and the accessibility of the VR experience, certain types of people (i.e., industry professionals, community members, students and researchers, figure 11) are more likely to play the game/try the experience, as they are more likely to visit Robohouse.

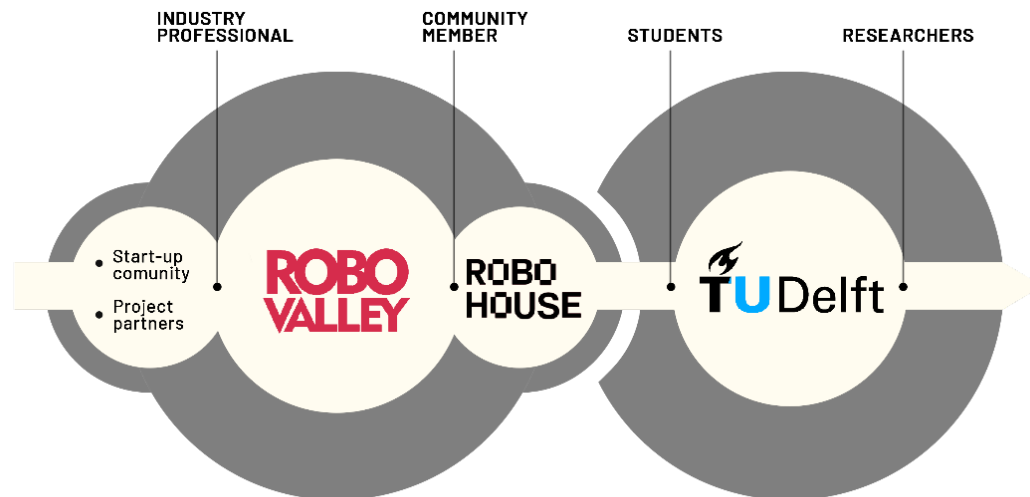


FIGURE 11. RELEVANT USERS OF THE EXPERIENCE

VR EXPERIENCE

SEQUENCE

The VR experience consists out of three distinct parts. In the pregame, players are eased into the game by getting to know the controls. This is done by answering a question through picking up and dropping a cube. The main sequence of the game contains the plausible future of work scenario of collaborating with different robots with the help of AR. Three wheels will be assembled to bicycle frames, carried and (re)placed by several robots, after putting on AR glasses at the end of the pregame (factory phase). The experience ends in the postgame after the third wheel is assembled, where the player will be transported back to the scene of the pregame.

Pregame

When the experience starts, the player is put in front of a big screen (number 1 in figure 12). Here, contextual objects, like bicycle frames, automated guided vehicles (AGVs) and boxes, are placed, outside of the play area of the player. When the player looks at the board, a brief explanation is given to the player. After the trigger is pressed, the player is asked a question about his/her 'work style'. The question can be answered by picking up a block and dropping it into a specific box, corresponding to a specific answer (2). After doing so and pressing the trigger again, the scene will change and the player will be put in the factory environment. This part is viewed as a tutorial by the developers.

As the player has just entered the factory, all the lights are still turned off, except for one spot light shining heavily on a pair of glasses (3). After the player grabs and puts on the glasses, it is revealed that the pair were AR glasses and an extra information layer is initiated in the view of the player (4). It welcomes the player and directs and guides him/her towards the lights to turn it on. The player then receives specifics about the tasks (5), whereafter the 'main sequence' of the experience starts.

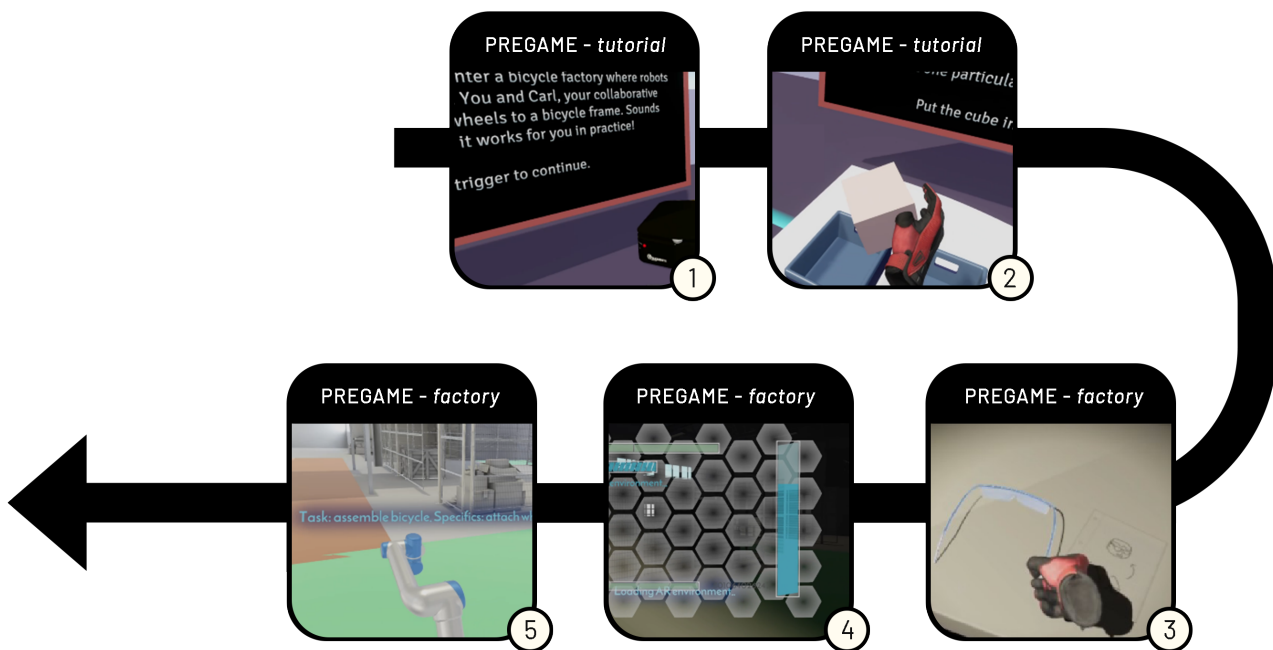


FIGURE 12. PREGAME FLOW

Main sequence

The main sequence starts with an AGV bringing a bicycle frame to the robot arm (figure 13, number 6). As the bicycle frame gets picked up and repositioned by the robot arm, an in-game arrow directs the player towards a rack of wheels on the left (7). A specific wheel needs to be picked up by the player (8) and placed onto the frame. After this has been done (9), the user is guided towards a drill (10). The drill is then used to attach the wheel onto the bicycle frame (by pressing the trigger on the highlighted spot, 11). After the user puts the drill back and spins the attached wheel, the robot arm puts the assembly back on the AGV (12 and 13 respectively). The AGV drives away and a new frame will be brought by another AGV, whereafter the sequence starts over again.

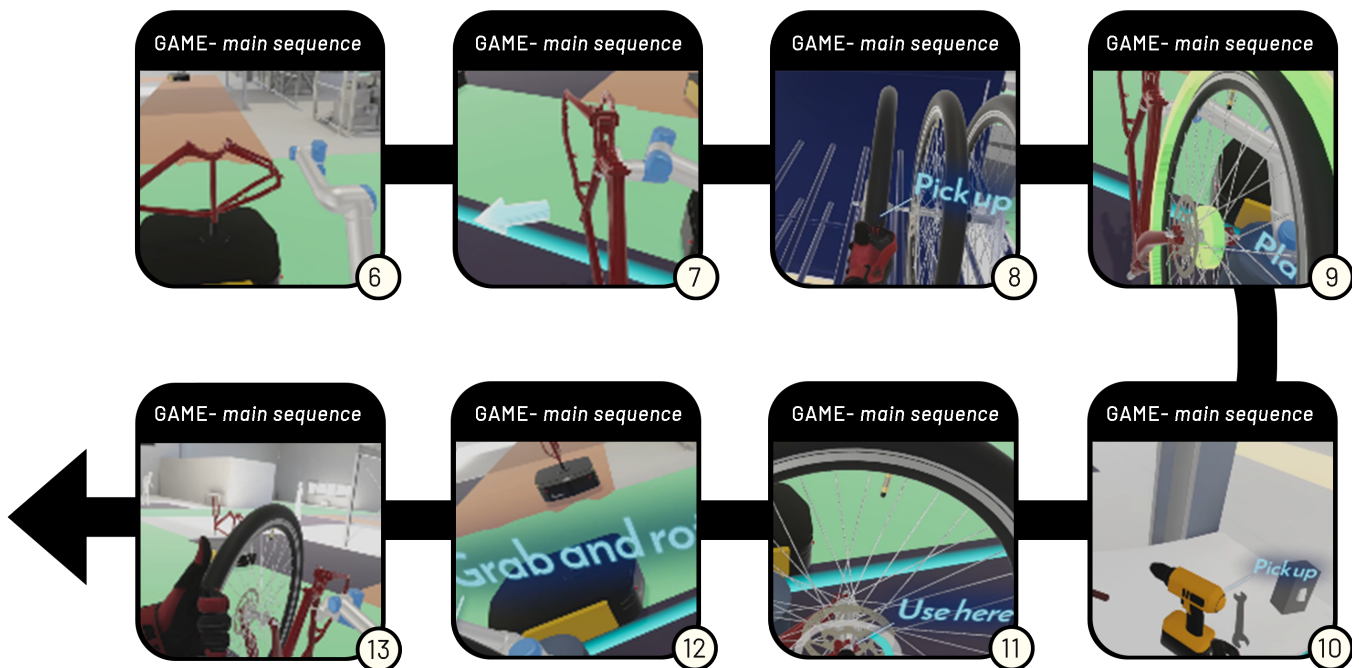


FIGURE 13. MAIN SEQUENCE FLOW

When the main sequence starts for the third time, an error will occur after a wheel is placed (figure 14, letter A). After analyzing for a brief moment, the AR layer will relay to the player that the wrong wheel has been placed (A) and the currently placed wheel needs to be put back into the rack again (B). The user is then informed that a new order has been placed. Two pairs of two AGVs are then seen driving towards the player. The first pair takes the present rack away, making place for the newly ordered rack of wheel, brought by the other pair of AGVs (C). When the rack is replaced, the user gets a new message to pick up and assemble a wheel. This time the last wheel in the rack needs to be assembled (D). After the player picks up and places the wheel (8 and 9 respectively), the sequence (depicted from numbers 10 till 13) continues. This will be the third and final frame assembly, including the error in this round, the player makes.

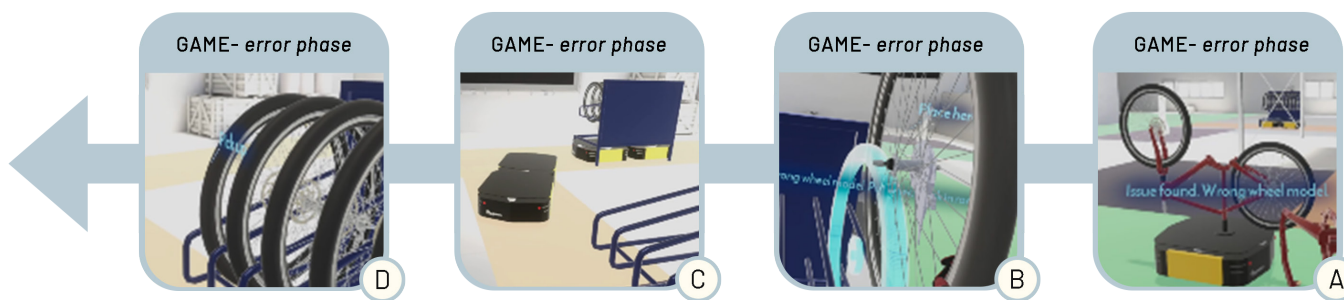


FIGURE 14. MAIN SEQUENCE FLOW, ERROR PHASE

Postgame

When the attached wheel is turned (14), the experience is over and the user receives a message to turn off the lights. The user is then directed towards the light switch (15). When the light switch is flicked, the user receives a message that the factory is being shut off. Finally, the player is transported back to the environment of the pregame, which marks the end of the experience.

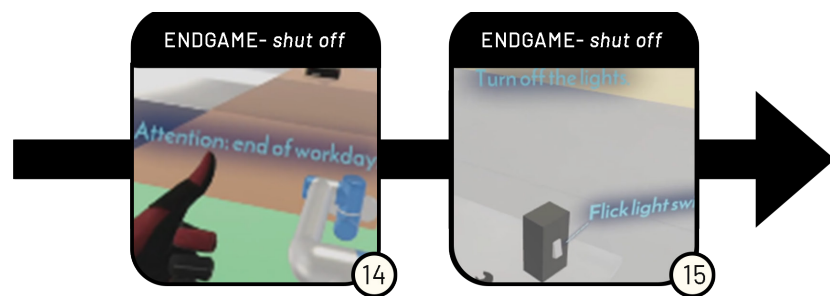


FIGURE 15. POSTGAME FLOW

Figure 16 shows the flow of the experience. Starting from 1, the player will go through the tutorial and assemble the first wheel to the bicycle frame (13) in the main sequence. The frames are carried by AGVs and lifted by the robot arm the player collaborates with. The second assembly is made following the same pattern, but starting from 6 instead of 1. When the third wheel is placed (9), an error sequence (A – D) will occur, where a new rack of wheels will be brought by AGVs after the player puts the already placed wheel back into the rack. From here, the last assembly (third) is made from 8 – 13 (skipping the letters). After, the postgame starts and the player gets transported back to the tutorial scene (1) following 14 and 15.

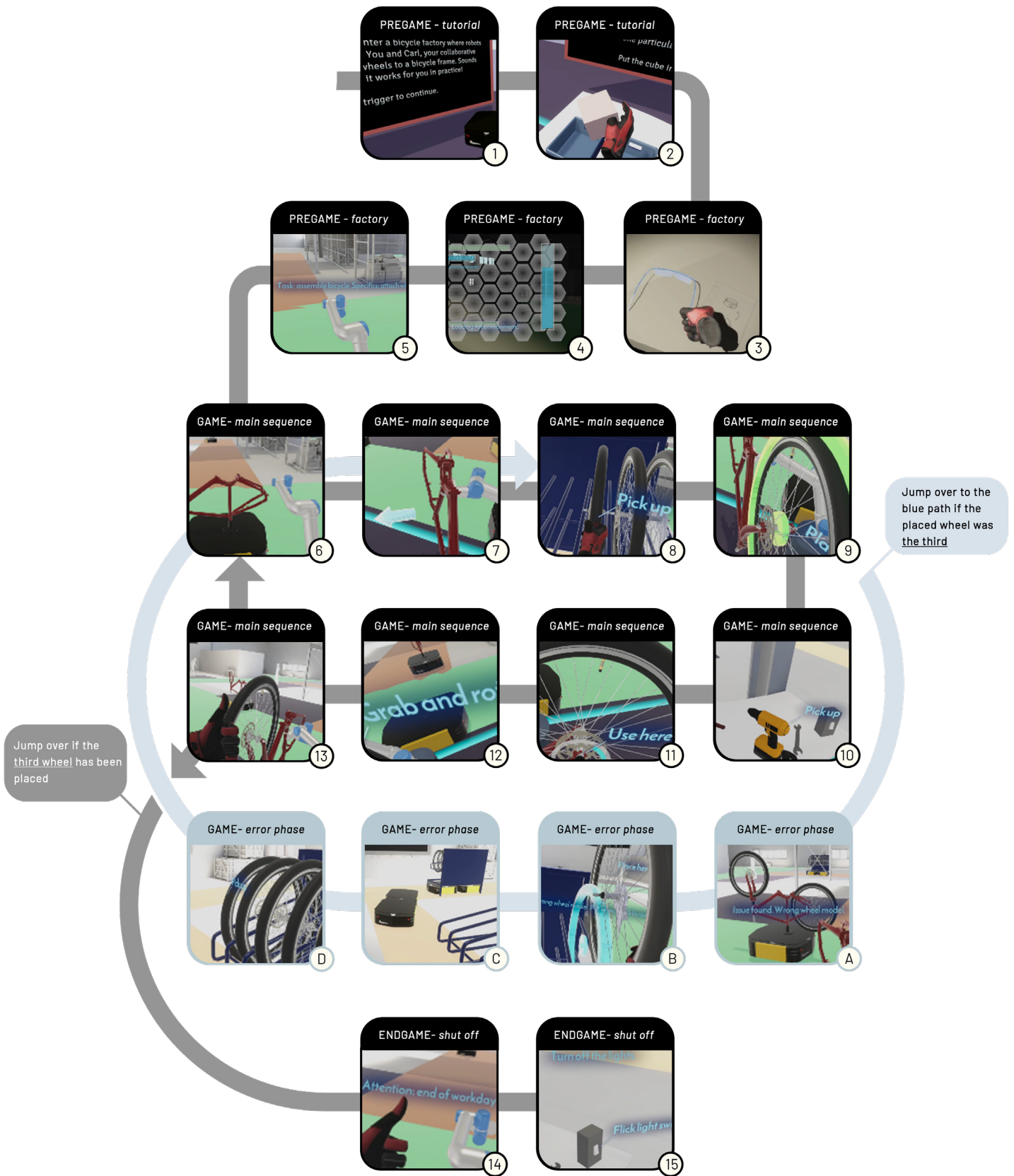


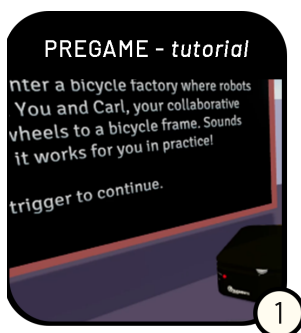
FIGURE 16. VR EXPERIENCE SEQUENCE

PRELIMINARY ANALYSIS

A preliminary test has been conducted before the start of this project with students of the TU Delft (a subset of the relevant users in figure 11). This test was focused on solving (relatively superficial) usability issues in the game and gauging the user-experience. Observations were used for the former, while an interview was used for the latter. Data collected (recordings of gameplay and interview) have been reviewed during this project. The interview setup, together with clusters made from the interview can be found in appendix A and B respectively.

OBSERVATIONS

Observing the 6 participants, 10 general errors have been identified. Eight errors are directly tied to the (in)ability to act in VR and occur at specific locations in the game (see flowchart in appendix D). These are:



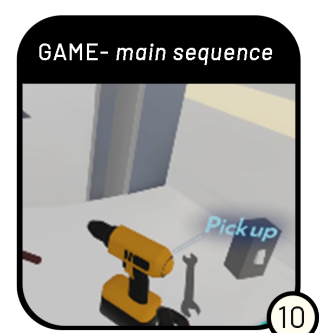
ERROR 1



ERROR 2



ERROR 3



ERROR 4

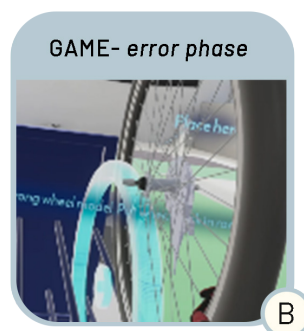
1. Using the controller to press the trigger in order to advance to the next scene
2. Using the controller to press the button in order to grab (or release) an object (block)
3. Using the controller to press the button to grab (and hold) the glasses to put on
4. Using the controller to press the button to grab (and hold) the drill.



ERROR 5



ERROR 6



ERROR 7



ERROR 8

5. Using the controller to press the button to activate the drill (while grabbing)
6. Interacting with a game object (grab and rotate the priorly attached wheel) in VR
7. Interacting with a game object (grab the priorly attached wheel) to put somewhere else
8. Identifying the object (grab last wheel on newly ordered rack) that needs interaction in VR

Participant	Problem number								Count	Pp	%	Playtime (in minutes)
	Pregame			Game/Main sequence			Error phase					
	1	2	3	4	5	6	7	8				
1		1	1		1		1	1	5	5/8	63	08:01
2	1		1		1	1		1	6	6/8	75	13:00+
3			1		1				2	2/8	25	05:27
4			1		1				2	2/8	25	08:11
5		1		1				1	3	3/8	38	07:13
6	1		1		1				3	3/8	38	07:38
Pe	2/6	2/6	5/6	1/6	5/6	1/6	1/6	3/6				
%	33	33	83	17	83	17	17	50				

TABLE 1

Column-wise, the chances (Pp) in table 1 represent the number of total identified errors found by each participant. When the time of completion of the game is combined with this, a general idea of the 'smoothness' of gameplay for the participants can be deduced; low number of errors encountered together with a fast gameplay-time generally equates to 'smooth' gameplay (e.g., participants 3 and 4), while a high number of errors together with a high game-play time would mean the contrary (e.g., participants 1 and 2).

Row-wise, the chances (Pe) represent how many times a specific problem has been encountered in total. It gives an impression of how frequently an error occurs. As the identified errors are location bound, higher frequency would equate to a higher number of participants experiencing the same problem at the same location.

Noticing and putting on the AR glasses are (error 3), together with activating the drill while holding it (error 5), usability issues that almost all participants encounter (83%). Also, half of the participants failed to notice which game object required action after the AGVs replaced the rack (during the error phase in table 1). Different variations of grabbing are troublesome for most participants. Although this basic interaction error occurs less frequently in a specific location when compared to the aforementioned errors, variations of basic interaction errors do occur often when counting multiple locations (errors 4, 6 and 7).

Two additional errors have been identified aside of the eight mentioned. They mainly revolve around participants not knowing what to do and not being immersed in general. These two errors are, as opposed to the other eight mentioned, observable throughout gameplay (not tied to one location in the game). These are:

9. Players tend to not move much within the play
→ E.g., players bend over in order to bridge the distance in front of them.
10. Players often interact with game objects before the game requires this to take place or the game doesn't require at all.
→ E.g., Wheel and other tools on the desk in the pregame – factory part (especially when the spotlight cue of the AR-glasses is overlooked).
→ E.g., Wheel in the rack before the robot arm has positioned the bicycle frame during assembly in the main sequence of the game.

FLOWCHART ANALYSIS

As part of the analysis of the game, a flowchart has been made; with every step the user and the system makes, combined with the observed errors, included. As the entire flowchart is relatively big and only needed during analysis, it has been omitted here, but added to appendix D. Highlights of the chart will be given whenever necessary.

Figure 17 is a simplified version of the entire flowchart. The figure combines the total of observed usability and clarity errors made, per stage of the game.

It is evident that there should be more focus on basic controller functions before entering the factory in the game; from using the trigger (e.g., to advance to the next scene), to using the controller to grab and interact with other game objects in the scene (e.g., activating a grabbed game object). Remarkably, not a single participant went through the experience without encountering usability issues. All of them had basic interaction troubles in some way or another.

Many errors already occur at the start of the game, during the tutorial phase of the pregame. As these errors are mostly left unresolved, once the player gets transported into the factory, more errors are encountered. Notably, during gameplay of the main sequence, the number of usability errors made in each round drop considerably. However, during the error phase, as new interactions are required, more issues are counted due the game being unclear. Furthermore, only during the second assembly in the game will some players experience the gameplay without usability errors. When the tutorial phase of the pregame is excluded (or the factory phase of the pregame is included), usability errors persist until at least half of the experience before improving (for some).

Additionally, some players have troubles getting immersed into VR in general; as they tend to keep the physical environment in mind. Ideally, players understand that within the bounds of the play area, walking physically (to bridge distance) is allowed.

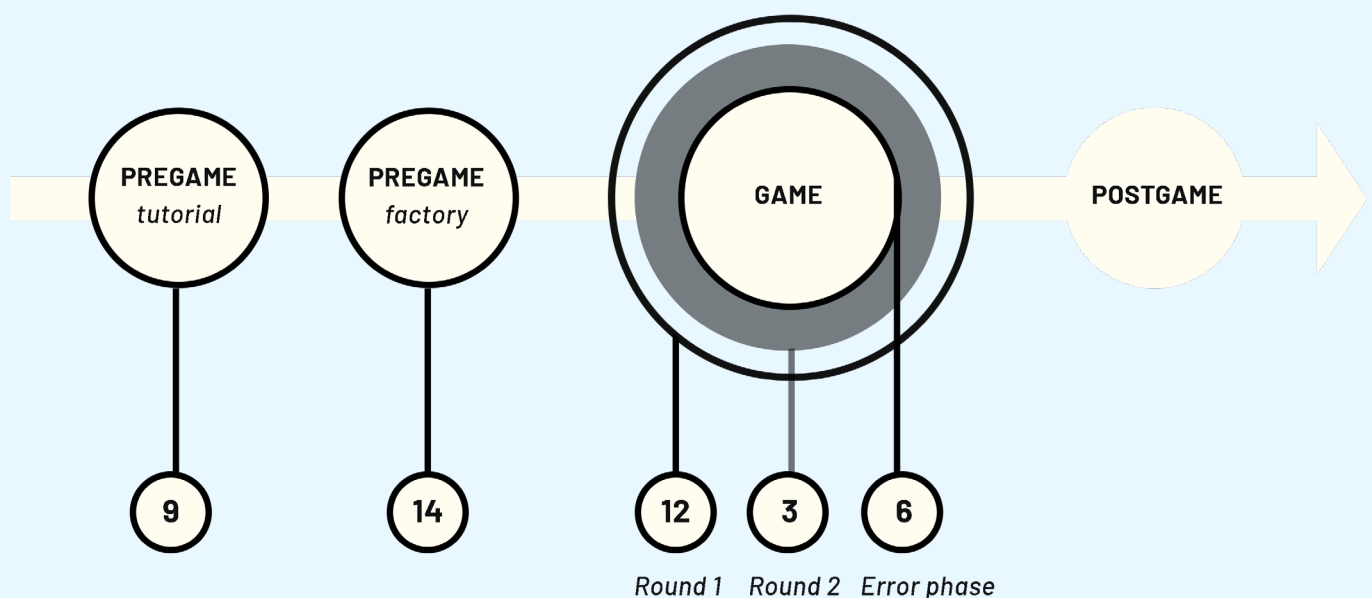


FIGURE 17. SIMPLIFIED FLOWCHART WITH CUMULATIVE NUMBER OF PRELIMINARY ISSUES

INTERVIEW ANALYSIS

Minutes have been taken from the recorded interview footage. The answers have consequently been translated and structured according to the questions asked in Miro. From this, (re)clusters have been made according to similarities and patterns found in the answers given by the participants.

The interview (and preliminary study in general) was relatively short, it consisted out of four questions. The first two inquired about the game itself, while the latter two gave participants an opportunity to propose for fixes. Naturally, a lot of varying answers were given for each question. However, in response to the question, relevant topics were pointed out by the participants, many of which were shared. Below, a visual is constructed by using the (usability and experience) topics participants shared, deduced from the first two questions. These will be connected to the fixes participants propose in question three.

“How did it go?”

Answers given by participants varied from ‘scary’ to ‘great’. However, they all mentioned the controls to be difficult in response to this question. All participants needed external help, given by the researcher, in order to overcome various (mostly basic interaction or clarity) issues. Half of the participants furthermore expressed the game to lack overall clarity (i.e., goals in the game, reasons for certain interactions).

“What was the impression you got (from the game)?”

Regarding the content of the game, most point out the austerity of the factory environment; specifying mostly the (missing) textures, sounds and people, and the general use of colors and lighting (4 out of 6 participants). Interestingly, half of the participants also noted the use of VR technology; generally stating the experience of it (i.e., VR in general as opposed to the content shown) to be very cool.

Fixes proposed (from question three) are connected to the relevant mentioned areas of improvements, making the following visual:

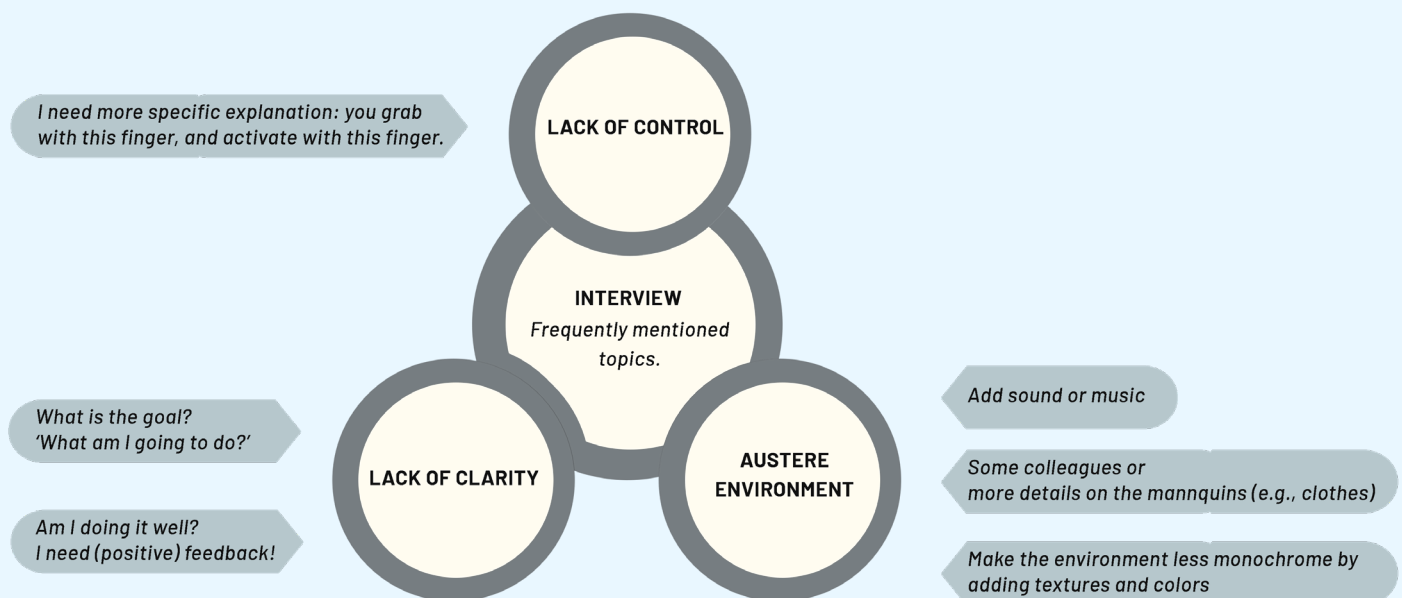


FIGURE 18. INTERVIEW TAKE-AWAYS WITH PROPOSED FIXES FROM PARTICIPANTS

While participants haven't directly stated to not be immersed, and this question was also not posed nor measured; troubles with immersion can be reasoned from the topics participants mention as response to the questions posed ("How did it go?" and "What was the impression you got").

In response, participants unanimously mentioned (the lack of) control, supporting the earlier observed need for more focus on basic controller functions in the (pre-)game.

Participants also mention a general lack of clarity. One participant expressed the experience to be confusing to the point that it made her feel insecure and scared. Additionally, others have expressed the need for affirmation or positive feedback also.

When talking about the environment, participants describe it generally as 'austere', mostly using words like cold, clinical, straight-to-the-point etc. When inquiring further, three subjects are mentioned: the lack of sound, the 'isolated' experience and the (clinically) clean environment.

The user experience of the content of the game can be very confusing. The game lacks overall clarity and the surroundings in VR leave a general austere impression on the participants. Half of the players found the game confusing and all of them mentioned usability issues, some even expressed feelings of insecurity. Notably and contrastingly, the user experience of the technology of the game (VR) is, even with the priorly mentioned (usability) issues, generally satisfying.

CONCLUSION

It is evident that the pregame (figure 12 images 1, 2 and 3) does not establish much for the players. Not only was this observed and unanimously mentioned during the interviews; all six participants had negative attitudes towards the pregame (most mentioned it to be lacking), it can also be deduced by the -initially high- but overall decreasing number of errors made during gameplay (figure 17). Because usability issues encountered (e.g., grabbing and triggering of the drill) are solved relatively quickly through trial and error in the two rounds of gameplay (in the main sequence), it can be concluded that the short tutorial phase of the pregame failed in preparing the users prior to entering the factory.

Furthermore, because players are generally lost without external instructions given by the researcher, which is a notion that is further substantiated in the interviews, the experience can be very confusing. Also, not all cues given in the game are understood by the player. For instance, although players notice the spotlights on the (AR) glasses in figure 12 (before starting the main sequence of the game), the focus is not noticed as a cue for interaction (i.e., for the player to grab the object in focus). Similarly, during the third assembly, after a new rack of wheels has been brought to the player by AGVs (figure 14), half of the participants failed to notice the cue attached to the last wheel in the rack (figure 14D).

Although the interview did not ask the players specifically, it is questionable whether the 'future scenario' of collaborating with robots with the help of AR is conveyed effectively (or at all).

Notably, even though players generally describe the content in the game as austere, the experience of VR was satisfying for most. Furthermore, the combination of not knowing what to do (or why) in general, together with not knowing how to do it, results in a subpar user- experience of the game.

The tutorial needs to be redeveloped in order to solve the aforementioned usability and clarity issues. This will improve the overall user experience of the VR game, making the future of work scenario much more immersive. In order to find state-of-the-art and relevant fixes, a literature study will be conducted next to search for criteria that makes VR experiences immersive and intuitive.



Literature study

Virtual Reality (VR) can be defined as “the sum of the hardware and software systems that seek to perfect an all-inclusive, immersive, sensory illusion of being present in another environment” (Biocca & Delaney, 1995). Many researchers agree that VR consists out of three core components: immersion, (tele) presence and interactivity (Ryan, 2015; Mütterlein, 2018; Radianti et al, 2020). This literature study aims to seek whether any ‘criteria or guidelines for making intuitive immersive VR experiences’ could be found or otherwise derived from those components. As preliminary testing has indicated usability and immersion issues with the game (table 1), the criteria or guidelines will be used to improve on the game using state-of-the-art information and insights. Outside of the scope of this project, it might also help designers and developers create user friendly VR experiences for head mounted devices (HMDs) in general.

RQ: “What are criteria for making intuitive and immersive Virtual Reality experiences?”

‘Intuitive’ can be described in many ways, but they all center around the user experience. Designing products centered around the (human) user experience has substantial economic benefits as highly usable systems and products tend to resonate more with (end) users. They, generally speaking, show to be more successful/less prone to failure after launch. According to ISO 9241-210, user experience (UX) is “a person’s perceptions and responses that result from the use or anticipated use of a product, system or service”.

One of the distinctive qualities of VR technologies is its potential to immerse its users. Although many would agree on immersion as ‘distinctive quality’, the same cannot be said about how people define immersion and VR or its components. Some have been researched thoroughly, also outside of the context of VR (e.g., immersion as flow in the context of gaming), but when comparing studies; component descriptions seem inconsistent at times (e.g., presence and immersion, as their concepts are closely related) or haven’t been described in the same manner (objective vs subjective immersion).

Because of this ongoing debate about VR and immersion, no clear criteria for making intuitive and immersive experiences for VR could be found.

The goal of this exploratory literature research therefore evolved from ‘finding criteria or guidelines for making intuitive immersive VR experiences’ to ‘creating criteria or guidelines for making intuitive immersive VR experiences (using state-of-the-art insights and information gained).

METHOD

In order to get the latest VR insights, a systematic approach for finding relevant papers (figure 18) has been set up according to the TUD guidelines (Technische Universiteit Delft, 2021). The main database used in the semi-automatic process is SCOPUS. However, other papers from referrals in papers or by researchers were included in the manual process. VR research is extensive and as the research continued, my understanding of the concept evolved. It also became apparent, as mentioned in the introduction, that comprehensive and consistent criteria could not be found.

Two samples of papers have been used to guide the literature research and answer the research question. The first sample, created from the semi-automatic and manual process, was used to get a state-of-the-art understanding of VR and its components and relations. From this sample, the research objective went from finding criteria (related to other research papers) to creating criteria (from an understanding of VR, derived from other research papers). Another sample was created through a manual selection in order to accomplish this.

SEMI-AUTOMATIC PROCESS

During the semi-automatic process, the scope, keywords and research questions underwent several iterations. Search concepts were derived from keywords in the research questions, which were refined using visualizations of the samples created. Visualizations were made using VOS viewer on the keywords in order to refine the concept with relevant synonyms and tailor the research question further. For instance, during query 3, keywords were derived from the question: "What are the criteria for a user-friendly VR experience?". This resulted in the search: (TITLE-ABS-KEY (guideline* OR factor* OR criteria)) AND (TITLE-ABS-KEY ((user OR human) PRE/0 friendly OR intuitive)) AND (TITLE-ABS-KEY (virtual PRE/3 reality OR vr)) AND (TITLE-ABS-KEY (experience* OR UX)), which was used in SCOPUS to find 14 papers. The exact approach and results are attached in appendix E.

MANUAL PROCESS

After revisiting the first query, all papers from the second till the fourth queries were aggregated to create a list of 59 potentially relevant papers. The number is different because I could not download some of the papers found. After a relevancy check, consisting out of reading the titles, abstract, introduction and conclusion; 7 papers were analyzed before the direction of the research changed slightly. After gaining a more profound understanding of VR due to the process, and not being able to find satisfying criteria and elements of intuitive VR experiences from the sample; five more papers were selected and analyzed. These papers were found from referrals in papers and from other researchers. A combination of papers from both the first and second sample were finally used.

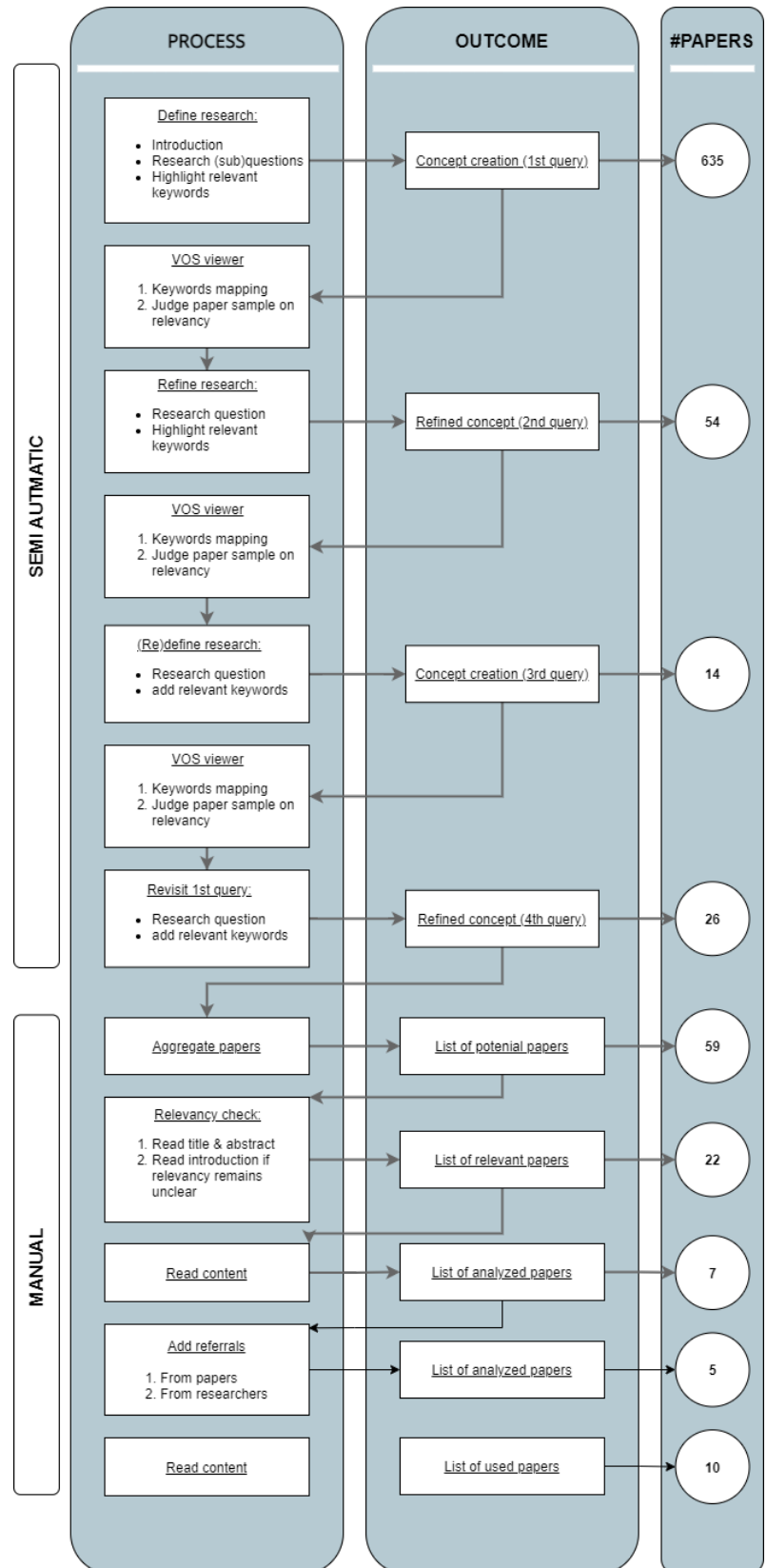


FIGURE 18. LITERATURE STUDY APPROACH

PRESENCE AND INTERACTIVITY

PRESENCE

According to Steuer (1992), presence is defined as “the extent to which one feels present in the mediated environment, rather than in the immediate physical environment” (telepresence refers to the state of the experience of presence that is reached using a medium). This phenomenon has been widely researched and several methods have been derived in order to evaluate presence. One of these is the presence questionnaire (PQ), which researchers like Joshi et al. (2021) have used to evaluate and improve VR experiences.



REALISM

3. How natural did your interactions with the environment seem?
4. How much did the visual aspects of the environment involve you?
- 5. How natural was the mechanism which controlled movement through the environment?**
6. How compelling was your sense of objects moving through space?
7. How much did your experiences in the virtual environment seem consistent with your real-world experiences?



POSSIBILITY
TO ACT

- 1. How much were you able to control events?**
2. How responsive was the environment to actions that you initiated (or performed)?
8. Were you able to anticipate what would happen next in response to the actions that you performed?
9. How completely were you able to actively survey or search the environment using vision?



SELF-EVALUATION
OF PERFORMANCE

15. How quickly did you adjust to the virtual environment experience?
16. How proficient in moving and interacting with the virtual environment did you feel at the end of the experience?



POSSIBILITY
TO EXAMINE

11. How closely were you able to examine objects?
12. How well could you examine objects from multiple viewpoints?
- 19. How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities?**



QUALITY OF
INTERFACES

14. How much delay did you experience between your actions and expected outcomes?
17. How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?
- 18. How much did the control devices interfere with the performance of assigned tasks or with other activities?**

PRESENCE QUESTIONNAIRE

The questionnaire consists out of a varying number of questions, asked for each of the five sub-scales: *realism* (questions 3 – 7, 10 and 13), *possibility to act* (questions 1, 2, 8 and 9), *quality of interface* (questions 14, 17 and 18), *possibility to examine* (questions 11, 12 and 19) and *self-evaluation of performance* (questions 15 and 16). When looking at the sub-scales and the questions of the questionnaire, it becomes clear how the degree of presence is evaluated and how it relates to VR experiences. The questions paint a picture of the sub-scales, which can be scored individually to reveal specific shortcomings regarding presence. The 19 questions of the PQ have been categorized into their sub-scales and can be seen on the left.

What stands out is how many questions are related to usability (specifically questions 1, 5, 18 and 19). The questions are spread across every sub-scale aside of self-evaluation of performance; which is arguably a mechanism supporting usability itself. Because usability affects the majority (if not all) of the sub-scales of presence, it can be concluded that, even within the scope of presence (and not user-friendly VR experiences), usability is a significant affector of VR experiences.

Cybersickness

Worth mentioning is that 'cybersickness' has been found to be a serious inhibitor of usability in VR. Cybersickness has many similarities with motion sickness in how it affects humans. According to theories concerning motion sickness (e.g., sensory conflict theory), cybersickness may occur because of conflicts between the visual and vestibular system, caused by the content (experienced through the medium) and the direct environment. For example, in many simulations, the visual system of the observer will receive information that suggests movement, while the vestibular system registers a stationary status. This discrepancy leads to symptoms of simulator sickness. In these cases, users experience for example headaches and nausea during or even after being immersed into a virtual environment (Stanney, 1995). Because of the similarity, the two often have been used interchangeably.

Following this similarity, it is argued that cybersickness is at its peak between the ages of 2 - 12. It then starts decreasing till around the age of 21, slowing down the pace till reaching non-existent levels at the age of 50. However, some report inconsistencies with this notion, as the contrary seems to be observed in virtual environments; where the elderly are more likely to get sick in VR than young players (Mousavi et al., 2013). Therefore, Arns and Cerney (2005) argue, the two concepts may not be similar at all, at least not concerning the relation of age with cybersickness and motion sickness.

Mousavi et al. mentions several other factors, aside of age, that "have shown considerable contribution to cybersickness in VE". These are:

- Unclassified factors. Display and technology issues (e.g., lag and flicker) and position tracking errors.
- Gender. It is argued that, because women have a wider field of view (FOV), they are more susceptible to flicker issues and generally are more prone to experiencing cybersickness.
- Illness. If a person has an illness, they may experience cybersickness differently.

It is difficult to draw conclusions on cybersickness based on any factors, as results are mixed and seem to vary much. Contributing factors that researchers point to (like the factors mentioned above) are therefore very generic and doesn't help much from a design-perspective. Generally speaking, VR experiences need to consider the possibility of cybersickness, especially if the experience is relatively long or meant to be played by relative older people.



FIGURE 19. OMNI ONE BY VIRTUIX

Simulator Sickness Questionnaire

The Simulator Sickness Questionnaire (SSQ) has first been proposed by Kennedy et al. (1993), almost three decades ago, and has since been used and peer reviewed widely (e.g., Balk et al., 2013). The questionnaire allows participants to score sixteen symptoms on a four-point scale (0 – 3); ranging from 'none' (0) to 'severe' (3). Four representative scores can then be calculated using a factor analysis: Nausea associated sub-score (N), Oculomotor associated sub-score (O), Disorientation associated sub-score (D) and the Total Score (TS), which is a weighted sum of the three sub-scores. The scores (N, O, D and TS) give the researcher an overall description of the simulation with regards to simulator sickness. When the SSQ was peer reviewed by Balk et al., they not only validated the questionnaire and its results, they also found that high SSQ scores, in particular a high sub-score for nausea, is more likely to cause (or in the words of Balk et al. "strongly correlated to") participant-dropout. Furthermore, scenarios that increase the possible discrepancy between visual input given by the medium and the participant's vestibular system, like stops, turns and curves, are also more likely to induce simulation sickness symptoms.



FIGURE 20. ENTRIUM 4D BY SAMSUNG

As mentioned in the previous paragraph, the SSQ is (like the PQ) very established. The questionnaire has therefore been used often when evaluating VR experiences. For instance, Liu and Uang (2007) used the SSQ in their research and concluded that prolonged exposure duration would lead to higher severity of cybersickness, especially for older people. Joshi et al. (2021) found, using the SSQ, that navigation velocity should be kept low in order to reduce cybersickness in general.

Several solutions have been produced and proposed to reduce or even eliminate cybersickness.

Technological solutions like motion platforms (e.g., the Omni One by Virtuix, figure 19) aim to solve the issue by allowing for physical movements. Other devices, like Samsung's Entrium 4D (figure 20), offer a solution through tricking the vestibular system. According to Mousavi et al. (2013), this is done via electrical signals that are sent to the 8th cranial nerve. This signal tricks the system into thinking that a linear or angular acceleration is taking place.

The most credited approach however, as argued by Mousavi et al. (2013) would be providing means for adaption. This allows people to be 'eased in' more softly and could also benefit general acceptance and usability.

INTERACTIVITY

According to Steuer (1995), interactivity means “the degree to which a user can modify the VR environment in real time”. This definition has similar (if not the same) meaning with “the possibility to act” and other sub-scales (possibility to examine and self-evaluation of performance to some extent) of the PQ as this allows, depending on the level of implementation, for a degree to which a user is able to modify the VR environment fundamentally. In order to get a sense of what interactive VR elements are mainly used and how they can possibly connect to the sub-scales of the PQ, regularly used elements in successful VR implementations are looked at closer.

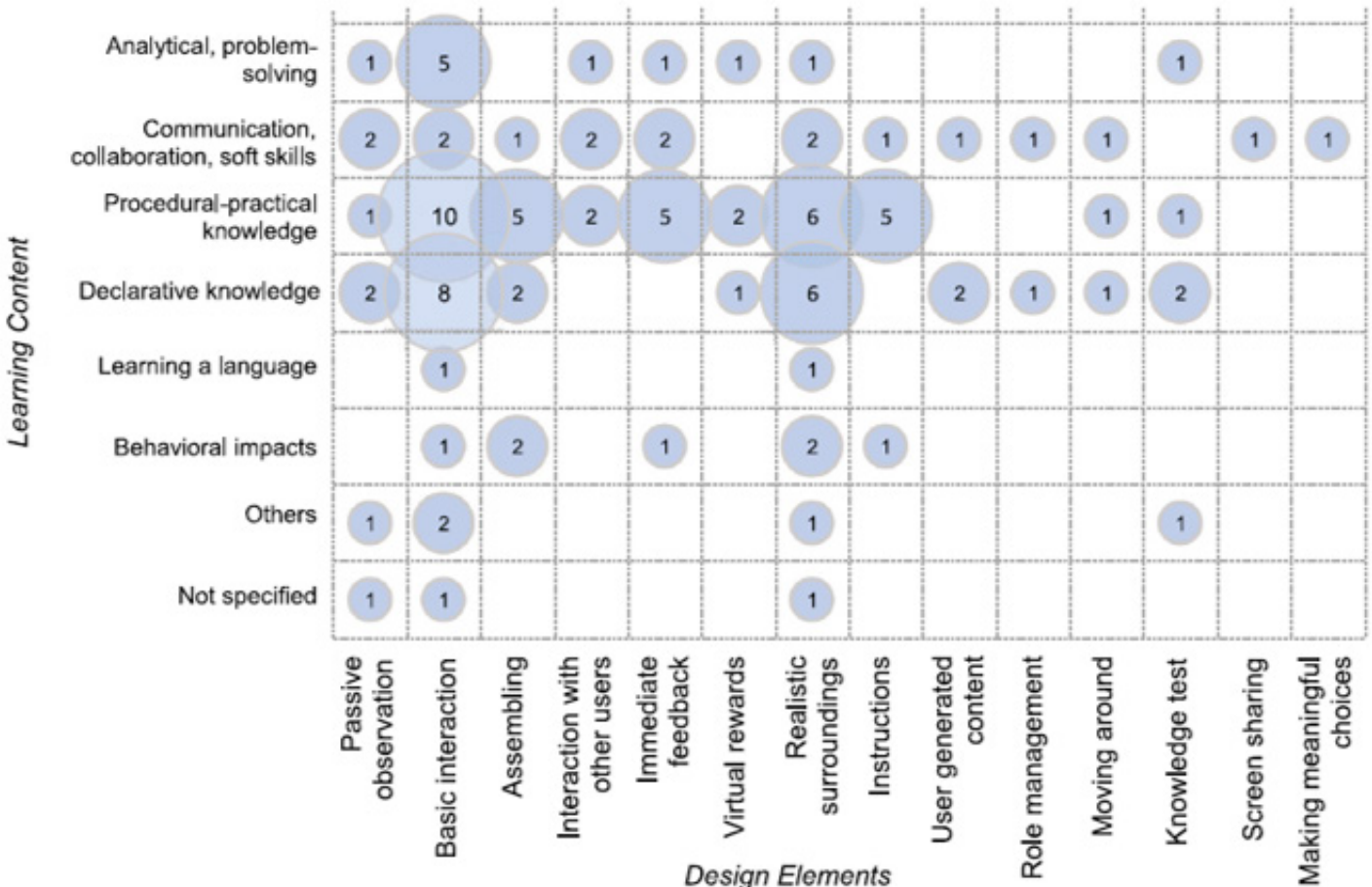


FIGURE 21. LEARNING CONTENT AND DESIGN ELEMENTS (RADIANTI ET AL., 2020)

Two systematic reviews regarding interactivity, as defined by Steuer, stood out because of their similarities and conclusions. Radianti et al. (2020) compared 38 successful implementations of VR in education (with varying didactic topics) and Chavez and Bayona (2018) looked at 30. While the two reviews had differences: Chavez and Bayona looked at education more broadly (also including training and the use of non-immersive VR), while Radianti et al. specifically focused on the teaching aspect of education, incorporating also different learning theories in their review and only used immersive VR; they both had a (part of their) conclusion dedicated to VR elements in the experience. This makes their outcomes, especially the elements mentioned by Radianti et al., extremely valuable when considering VR elements and (the sub-scales of) presence.

It becomes clear which VR elements are implemented the most (in successful VR experiences) when looking at the figure Radianti et al. produced (figure 21). In this figure, VR elements used are mapped against the learning content of that experience. Although the order of frequently used elements shifts a little bit depending on the learning content (e.g., when procedural knowledge is taught, the ability for assembly jumps up from the fifth most used VR element to the third), generally speaking, “basic interaction, assembly, passive observation, immediate feedback and instructions” are VR elements that are present in most successful implementations.



BASIC INTERACTION

Radianti et al. (2020) found that basic interaction and realistic surroundings were the most frequently used VR design elements in the articles they reviewed (24% and 17% respectively). Basic interaction was the most frequently used element and means the ability to look around, but also the ability to manipulate objects; this includes rotation, grabbing and placement, and movement (locomotion).



REALISTIC SURROUNDINGS

There seems to be a bit of ambiguity when it comes to realistic surroundings (second place), as this is heavily dependent on not only the hardware used (Radianti et al. reviewed immersive VR), but also the person perceiving it. Generally speaking, in the context of immersive VR, the representation of the surroundings should be deemed 'realistic enough' (Radianti et al., 2020).



IMMEDIATE FEEDBACK

Immediate feedback and instructions share a third place (each accounting for 10% of the articles). Immediate feedback is a large element and includes different alert systems (e.g., sounds to signal success and failure), virtual hands and voices, the display of data(stream) that shows or updates the status of an object, tactile feedback, signs and highlights.



INSTRUCTIONS

Instructions show information but also include pause, repeat and recall features, preferably in an intuitive menu (e.g., where you can drag and drop). The ability to interact with different language modules also belongs to the instructions element.



INTERACTION WITH OTHER USERS

Interaction with other users were implemented in 9% of the articles Radianti et al. reviewed, being the fourth most used VR element.



PASSIVE OBSERVATION

Passive observation and assembly shared a fifth place, each element being present in 8% of the articles. Passive observation can be viewed as a more dialed back version of basic interaction; where they both share the ability to look around, but passive observation includes the ability to observe objects and watch (pre)recorded content and excludes movement and the ability to manipulate objects.



ABILITY FOR ASSEMBLY

The ability for assembly includes picking up and using tools (to manipulate objects) and collect and classify items. As opposed to passive observation, assembly can be viewed as a more advanced version of basic interaction.

The least used VR elements in the articles Radianti et al reviewed were: ability to move around (sixth place at 4%), ability for user to generate content (seventh place at 3%), having virtual rewards, role management and a knowledge test (all three sharing the eighth place at 2%), and the ability to screen-share and 'make meaningful choices' (ninth place at 1% each).

As mentioned earlier, Chavez and Bayona conducted a similar review where they also incorporated ‘non immersive VR’ (VE). In their conclusions, a table was produced with elements that were most frequently used. However, because of this inclusion, the list produced by Chavez and Bayona mainly serve as a support for the findings of Radianti et al. Notable and frequent elements identified by Chavez and Bayona will be included, the five most frequent elements are highlighted in figure 22, but others are omitted.

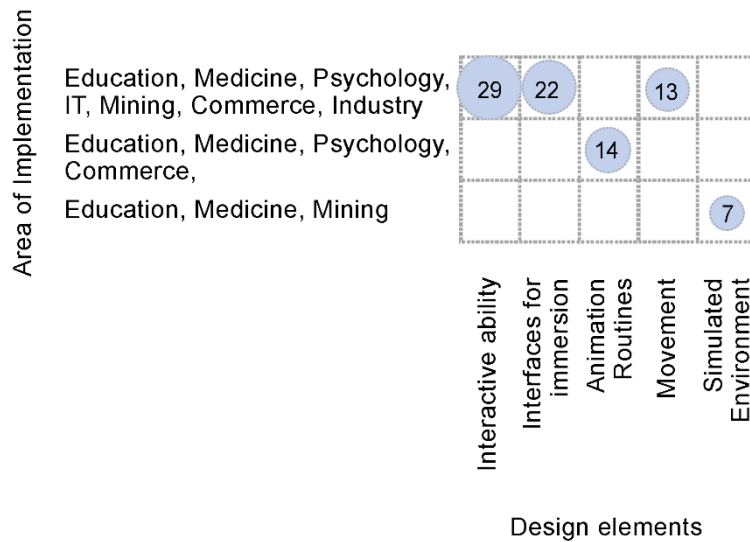


FIGURE 22. AREA OF IMPLEMENTATION AND DESIGN ELEMENTS (BASED ON CHAVEZ AND BAYONA, 2018)

The ability of players to interact with objects and the environment was the most important element, the authors deduced. In the articles they reviewed, they found that interactive ability was mentioned most frequently. Tools added to increase immersion was the second most important element. Chavez and Bayona also found frequent use of real scenes (to increase immersion) and animation routines (third place). Movement was the fourth most implemented element in their review. Interestingly, Chavez and Bayona noted the elements added in recent years, mainly contribute to enhancing the gamefulness of VR (narration, levels with timers).

Chavez and Bayona ranked different elements on the third and fourth place (animation routines and movement respectively) than Radianti et al. (immediate feedback and instructions respectively), presumably because they also included VE in their study. This could also help explain why simulated environment is ranked significantly lower (from second place in VR, to the fifth place in VR & VE), as the

environment becomes more important based on the immersive potential of the technology, which is much higher for VR than VE.

Using immersive VR technology also needs more focus on instruction and immediate feedback due to the difference in hardware used (and adopted), which is why the elements are much more prevalent in the review by Radianti et al. compared to Chavez and Bayona. This notion is further supported when viewing interactivity as ‘perceived interactivity’ (as mentioned by Mütterlein, 2018); where usability is the bottleneck before interactivity can take place (the VR environment can only be modified by the user when the user perceives it as modifiable). Having elements of feedback and instruction supports “the degree to which the user is able to manipulate the VR environment in real-time” (interactivity); which can be the crucial bridge in immersive VR.

CONCLUSION

When talking about presence “the extent to which one feels present in the mediated environment, rather than in the immediate physical environment”, usability was always of concern. Not only because (symptoms of) simulator sickness are/is so easy to appear in VR environments; being a serious inhibitor of overall usability, but also because when measuring presence through the PQ, all sub-scales are (somehow) related to usability.

1. **Possibility to act** - How much were you able to control events?
5. **Realism** - How natural was the mechanism which controlled movement through the environment?
18. **Quality of interfaces** - How much did the control devices interfere with the performance of assigned tasks or with other activities?
19. **Possibility to examine** - How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities?
0. **Self-evaluation** itself is a scale to support usability

Evaluation

The System Usability Scale (SUS) is a ‘quick and dirty’ evaluation tool, used across industries, that has shown promise regarding usability evaluations of VR experiences (Joshi et al., 2021). The survey is relatively short, consisting out of ten questions which you can answer by scoring on a 5-point scale. A total score can then be calculated. Although the score will not tell you what problem there is specifically, it will give you an indication of the usability of your system in general. The score can furthermore be compared to SUS results from other products.

Using this score in conjunction with the PQ to evaluate VR experiences can give the designer insight into usability and the overall quality. While the PQ itself takes usability into account, it doesn’t say anything about usability, but rather usability’s influence on presence (i.e., bad usability will have considerable impact on the evaluation of presence). This can be mitigated by involving the SUS.

Simulator sickness (or symptoms of it) are/is likely to happen when experiencing VR due to the technology used; at least, reasoned from a sensory conflict theory point of view: discrepancies between the visual and vestibular system, caused by the content viewed through the HMD and the direct environment respectively, will lead to (symptoms of) simulator sickness. It has shown to be a considerable inhibitor of general usability in VR. The Simulator Sickness Questionnaire (SSQ), like the SUS, has shown promise with regards to the evaluation of simulator sickness. Unlike the SUS however, the SSQ uses sub-scores to evaluate the overall experience of simulation sickness; giving insights into the specific contributions of the sub-items.

Using the SSQ, studies have shown that for VR experiences, nausea specifically should be avoided. Furthermore, increasing rotations speeds, navigation velocity and exposure duration will lead to a higher severity of simulator sickness experienced, especially for relatively ‘older’ people. Different solutions have been proposed, but providing for opportunities to ‘ease’ the players into the experience is, according to Mousavi et al., most effective.

Generally speaking; usability needs to be taken into account more explicitly when evaluating VR, as bad usability will have considerable effects on the overall experience. By making use of the PQ, in conjunction with the SUS and/or the SSQ, insights can be derived with regards to how immersive and intuitive (usable, not hindered by symptoms of simulator sickness) the experience is.

Sub-scales of the PQ and frequently implemented VR elements

The overlap in meaning of interactivity (as defined by Steuer; the degree to which a user is able to manipulate the VR environment in real-time) and ‘possibility to act’ (stated in the PQ) suggests a connection between the two. Therefore, the frequently used interactive elements in successful VR implementations (mainly found by Radianti et al.) have been connected to the sub-scales of the PQ in figure 23.

On the left, the definition of presence (what is measured using the PQ, as defined by Steuer), is shown. The sub-scales of the questionnaire are also visualized, together with its relevant questions.

On the right, elements for interactivity are summarized. As mentioned earlier, interactivity overlaps much with the possibility to act in meaning, but when looking at the groups of elements individually (black and white icons, e.g., basic interaction), more detailed connections between the sub-scales of the PQ and the elements of interactivity can be deduced.

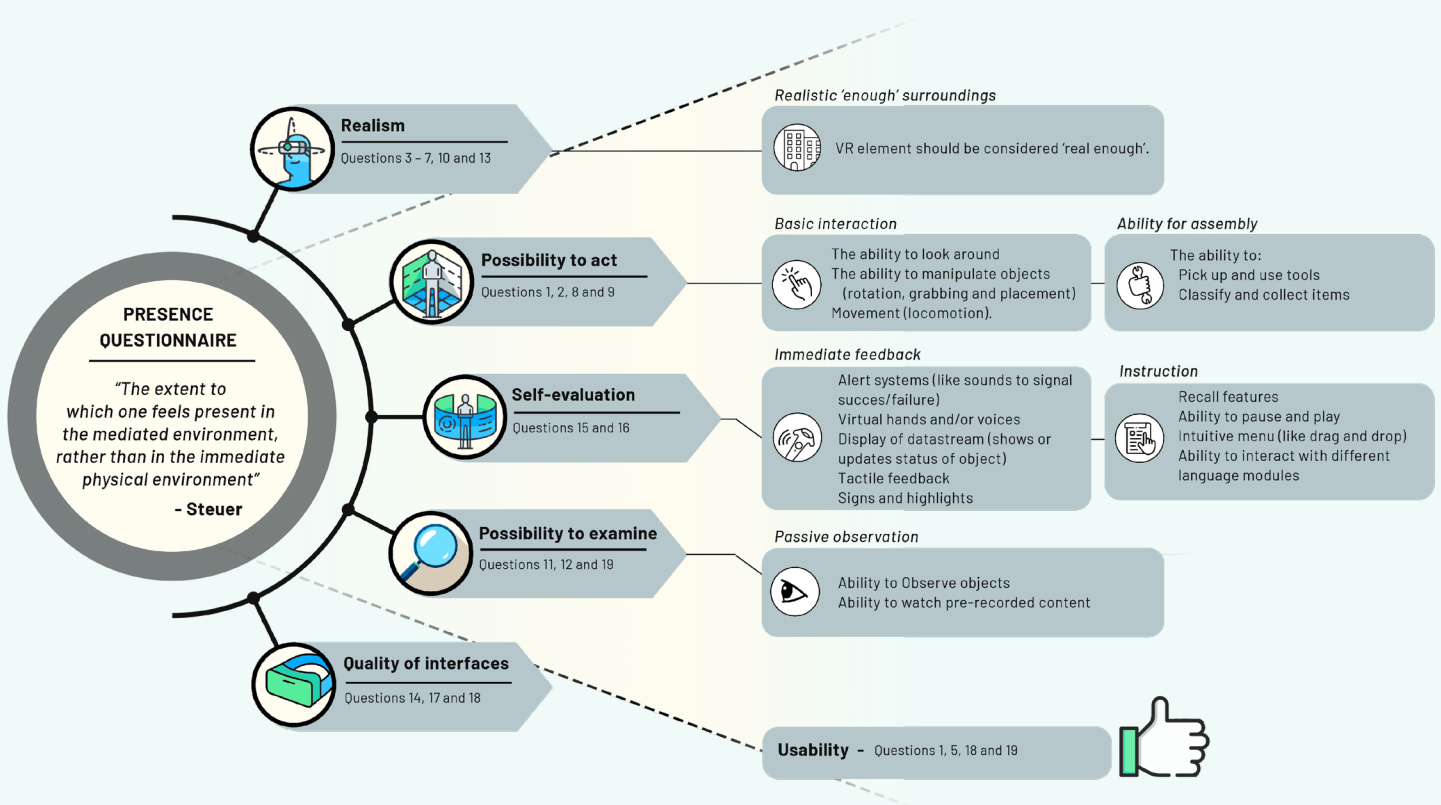


FIGURE 23. SUB SCALES OF PQ CONNECTED TO THE RELEVANT ELEMENTS OF INTERACTIVITY

Usability is related to interactivity (at least) through definition; as Interactivity, the degree to which a user is able to modify the VR environment in real-time, can only take place when the user perceives the VR environment as modifiable (or usable). How usability is related to presence has been elaborated on earlier. Because usability is prevalent in both presence and interactivity, it has been visualized behind both components as a funnel from the interactive elements to the related sub-scale of presence. The lines converge from interactivity towards presence (right towards left), as usability can be a bottleneck for interactivity to take place (and so also its influence on the degree of presence experienced).

Evaluating VR experiences using presence is a well-established method. This visual summarizes the insight of how presence, measured through the PQ, can be improved by giving examples of frequently used interactive VR elements can influence a specific sub-scale presence.

IMMERSION

Immersion is a topic that comes up in the majority of VR researches. Although there seems to be somewhat of a consensus in defining presence and interactivity, defining immersion sparks more debate.

From a technological point of view, the degree of immersion can be attributed to specific hardware components and/or features. The frame rate, processing power to produce a no lag experience, display resolution and also potential external peripheral controllers can all be indicators of the degree of immersion in the produced experience. Immersion is described here as the “extent to which the computer displays are capable of delivering an inclusive, extensive, surrounding, and vivid illusion of reality” (Slater & Wilbur, 1997). Radianti et al. approached immersion in VR experiences using this definition.

From a psychological point of view, the degree of immersion varies from person to person and is therefore a subjective experience. Mütterlein (2018) supports this subjective angle and defines it by borrowing from insights from psychology on the nature of ‘flow’. Following these insights; immersion is described as “a state of optimal experience where one is completely absorbed and immersed in an activity”. An optimal experience being one where “a person is totally engaged and pre-occupied by a task, to the extent that not only the performance of an activity syncs with the person, but the person also loses his/her self-awareness and track of time” (Nah et al, 2014).

Although the debate on immersion is still ongoing, some researchers have used immersion, defined from the psychological point of view, to evaluate VR experiences (as opposed to using presence as earlier described).

IMMERSION AND VIRTUAL REALITY

Although the phenomenon immersion is relatively well-documented, difference in defining immersion makes it difficult to be used consistently to evaluate VR experiences with. At least, not in the way evaluating VR experiences by presence is using the PQ. Mütterlein however, underlined that the earlier mentioned components (interactivity and presence), need to be considered together with immersion in order to comprehensively explore VR, as the synergy of these three components makes it as amazing as it is.

By making use of questionnaire scales borrowed from the flow theory (figure 24), as mentioned earlier, Mütterlein uses immersion/flow, together with (tele)presence and interactivity to evaluate (satisfaction in) VR experiences using his own questionnaire. The technological point of view Radiant et al. took was mentioned by Mütterlein. He regards the technological angle ‘merely’ as a restriction of the capabilities of a system, which should not be used to solely define immersion.

Construct	
<i>Telepresence</i>	
TP1	The VR content created a new world for me, and this new world suddenly disappeared when the VR content ended.
TP2	When the VR content ended, I felt as if I returned to the "real world" after a journey.
TP3	I forgot about my immediate surroundings when I was using the VR content.
TP4	The VR content seemed to be "somewhere I visited" rather than "something I saw".
<i>Immersion – Fluency</i>	
FL1	I had no difficulty concentrating.
FL2	My mind was completely clear.
FL3	The right thoughts and movements occurred of their own accord.
<i>Immersion – Absorption</i>	
AB1	I didn't notice time passing.
AB2	I was totally absorbed in what I was doing.
AB3	I was completely lost in thought.
<i>Interactivity</i>	
IT1	The VR content allowed me to interact with the virtual world.
IT2	I had the feeling that I could influence the virtual world of the VR content.
IT3	The VR content was interactive.
<i>Satisfaction</i>	
SA1	Very dissatisfied—Very satisfied
SA2	Very displeased—Very pleased
SA3	Very frustrated—Very contented

FIGURE 24. CONSTRUCT OF MÜTTERLEIN'S INTERVIEW (2018)

Interactivity was regarded, like earlier studies, as the degree to which the environment was modifiable by the user in real-time. (Tele)Presence and immersion have some commonalities, which leads to confusion to some, and cases where the two are used interchangeably. They are however, distinctively different; as being present refers to “being immersed in a particular environment, even when the person is physically situated in another” while being immersed means “being immersed with performing an activity”. According to Mütterlein (2018), interactivity was found to have a direct and positive effect on immersion, but it also influenced the degree of telepresence participants experienced. Because interactivity has a direct and positive effect on all components, it is regarded as a foundational component of immersion by Mütterlein. Telepresence also had a direct and positive effect on immersion. Mütterlein further noted that a high level of immersion was directly and positively related to the satisfaction participants experienced while in VR (figure 25).

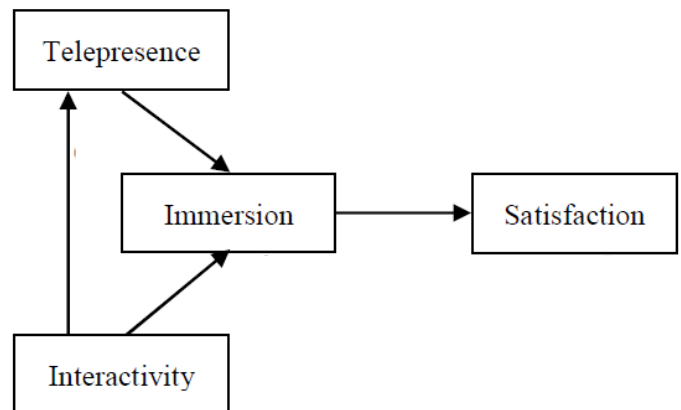


FIGURE 25. RELATION OF INTERACTIVITY, PRESENCE AND IMMERSION (MÜTTERLEIN, 2018)

FLOW

The experience of a deep level of engagement is what is referred to as flow. The experience takes place in a wide variety of activities, including participating in art and sports, playing a (video)game or performing surgery for instance. Because Mütterlein used flow in order to (help) define immersion, virtual reality is grouped among these activities implicitly. When flow is experienced, the individual enjoys the activity immensely and also loses track of time (Nah et al., 2014). It is an important theory when it comes to understanding engagement.

A systematic review by Nah et al. (2014) has researched this topic in the context of gaming (in general). The review seeks to suggest for game design elements that may be useful in order to induce flow; for a highly engaging, high involvement and immersive game.

Conditions of flow

Csikszentmihalyi (1990) explained that three (key) conditions should be met in order to experience flow: clear goals, unambiguous feedback and a balance of challenges and skill. Five conditions were later added (1997). According to Csikszentmihalyi, in the description of flow, people described their experience to be comprised of one out of eight conditions (at least).

Different researchers have used these conditions found by Csikszentmihalyi or an adjusted or subset version of it in order to measure flow (Fang et al., 2013 and Mütterlein respectively). Because there is no agreed standard on the dimensions of flow, Nah et al. created their own set by borrowing from the adjusted conditions of Fang et al. and adding one extra condition: enjoyment.

The four original and unchanged conditions (Csikszentmihalyi, 1997) are:

1. Clear goals. No matter what the task is, having explicit goals is needed for flow to take place.
2. Immediate feedback. This condition supports challenging activity that requires skill and clear goals as an immediate and clear feedback mechanism allows for a person to evaluate and improve towards the goal/task.
3. Concentration (on task at hand). When a person is totally engaged in a task and is experiencing flow, (other) thoughts are removed and distractions are ignored when fulfilling the task.
4. (Paradox of) Control. The person experiencing flow perceives to be in complete control of his/her actions. They believe that no mistakes are possible and may therefore participate harmful activities (if a mistake happens).

Challenging activity that requires skill (first condition that Csikszentmihalyi mentions) was altered slightly by Fang et al. into:

5. Balance of challenge and skill. These two need to be in balance where boredom does not take place because the skills are higher than the posed challenge, and anxiety doesn't occur because the skills are inadequate or the challenge is too demanding.

Nah et al. (2014) added another condition:

6. Enjoyment. This is a proven affectional outcome that arises from experiencing flow (Nah et al., 2014). It can be related to perceived feelings of competence also.

Furthermore, by using various validation and cleansing methods, three conditions of Csikszentmihalyi's measurement of flow were compiled into one by Fang et al.:

- Merging of action and awareness. When a person is totally engaged in the task, to the extent that the performance syncs with the person (as performing will become automatic and spontaneous), they are experiencing flow.
- Loss of self-consciousness. When flow is experienced, a person loses his/her self-awareness because they are being pre-occupied with an activity.
- Transformation of time: As stated earlier, when flow is experienced, a person's perception of time is distorted.

creating:

7. Immersion. When a person is totally engaged and pre-occupied by a task, to the extent that not only the performance of an activity syncs with the person, but the person also loses his/her self-awareness and track of time.

Game elements and Flow conditions

Nah et al. furthermore identified a set of game design element that induce flow during gameplay from mappings performed by Sweetser and Wyeth (2005) and Cowley et al. (2008). Combining these game elements with the conditions of flow (vertically in figure 26), Nah et al. explains how different elements affect flow. These explanations often go along with study conclusions from other researchers.

Flow component	Game design element					
	Storyline	Interactivity	PBL	Levels	Controller type	Social interaction
Balance of challenge and skill		X	X	X	X	X
Clear goals	X		X	X		
Immediate feedback		X	X			X
Concentration		X				
Control		X			X	
Enjoyment	X		X		X	X
Immersion	X	X			X	

FIGURE 26. MAPPING OF GAME DESIGN ELEMENTS TO THE FLOW CONDITIONS (NAH ET AL., 2014)

“When a person is totally engaged and pre-occupied by a task, to the extent that not only the performance of an activity syncs with the person, but the person also loses his/her self-awareness and track of time.”
 - Fang et al. (2013)

“a state of optimal experience where one is completely absorbed and immersed in an activity”
 - Mütterlein (2018)

Because Mütterlein used the theory of flow to define immersion, it's no surprise that his definition overlaps much with how Fang et al. defined it (and how Nah et al. used it). An “optimal experience” could even be interpreted as the state Fang et al. described. Interestingly however, is how immersion can be mapped as a dimension of flow itself. Nah et al identified three game elements influencing immersion (as dimension of flow); interactivity, controller type and storyline.



INTERACTIVITY

Interactivity, as game element, impacts the experience of flow the most. This is evident as it covers five out of seven conditions of flow, including immersion. This relation was also found by Mütterlein (2018) for VR, where he noted interactivity to be a fundamental component (of the interactivity, presence and immersion trio).

Nah et al. describe interactivity as ‘intuitive game mechanics’ which helps bring along the mechanics of a game. Intuitive interactivity contributes to flow by helping the balance of challenge and skill. With feedback, the user is able to concentrate and improve on a task, elevating not only (the feeling of) control, but also immersion.

The degree of control and concentration can also be influenced by features like informational cues, to help players balance the challenges posed by the game (reasoned from another perspective). Additional features and forms of feedback can help to deepen flow through its conditions (Nah et al., 2014).



CONTROLLER TYPES

Controller Types are (the interaction) medium between the players and the game. They affect the flow greatly (second after interactivity) through the conditions: balance of challenge and skill, control, enjoyment and immersion (Nah et al., 2014). Because of its great impact on control; interactivity and controller type are the only elements that influence the control dimension, poorly designed or improper use of controllers will affect flow greatly. Distractive activities like technology (e.g., connection and interface) issues have been found by Admiraal et al. (2010) to be negatively correlated with the experience of flow also. Therefore, tools used or information (cues) and objects shown in the game should not be misleading as this could disrupt the players experience of flow.



STORYLINE

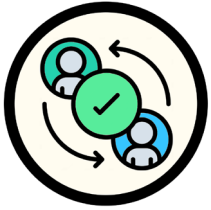
Storyline contributes to flow by addressing the conditions: clear goals, enjoyment and immersion. Storylines and narration help the comprehension of (high-level) goals of the game that gives meaning. Nah et al (2014) found that players who understand/relate to the storyline are able to perceive more information cues. Additionally, Sweetser and Johnson (2004) found that great storylines also enable higher empathy with game characters. This furthers the immersion and enjoyment conditions of flow.

Three game design elements are also noted by Nah et al. These elements don't (directly) influence immersion, but does help elevate the experience of flow (essentially through gamification)



POINTS, BADGES & LEADERBOARDS

Sharing the second place with controller type are Points, Badges and Leaderboards (PBL), which also influences enjoyment using elements to increase gamefulness. They also provide immediate feedback and can help provide clear goals (in order to experience flow) together with the element storyline.



SOCIAL INTERACTION

Social interaction is, together with the dimensions storyline and PBL, the main source for enjoyment (assuming proper usability of the controller type condition). Social interaction includes cooperation and competition between players (elevating the dimensions balance of challenge and skill and immediate feedback aside of enjoyment), but also more general interactions like messaging each other and sharing data on bulletin boards (for example).



LEVELS

The game element that covers the least number of conditions are levels (2). Arguably, it's the least important element, as the conditions addressed are not exclusive to the element (other elements, like PBL, also addresses the conditions levels cover, among others) in addition to relatively low contribution to (the conditions of) flow overall.

CONCLUSION

When viewing VR through experienced immersion (as opposed to experienced presence), viewed from the psychological point of view, interesting tangent areas arise. Flow has been, like presence, a very well-researched phenomenon. Defining immersion through flow (or as a dimension of flow) connects the topics and allows for interesting suggestions. In this chapter the connection of immersion and flow with game elements have been reviewed, mainly using the study of Nah et al.

Generally, all elements induce flow by tackling one or more of its conditions. Some seem more fundamental, like interactivity and controller type, compared to others (e.g., levels). This distinction in number of conditions met and impact made, makes it helpful in deciding which (combination of) element(s) fits what need and implementation. In figures 27 and 28, the game elements (targeting immersion and flow) have been ordered according to the number of flow dimensions they target (which will be mentioned first).

Immersion and flow elements

Three game elements influence immersion when viewing it as a dimension of flow. Like Mütterlein concluded, immersion is heavily influenced by interactivity. It targets control, enjoyment and immersion among others, but most importantly manage the balance of challenge and skill. Notably, 'intuitive' was mentioned often when talking about interactivity; i.e., elements should not only provide for interactivity, but also support and communicate it effectively.

Controller type is the second most influential element; it is the medium between the player and the game. Because of this, poor use or distractive issues (like connection or interface problems) will affect immersion greatly. Controller type includes (other) hardware used, this satisfies Mütterlein's opinion that the technological point of view of immersion (defined by Slater and Wilbur, used by Radianti et al.) "merely serve as a restriction" for immersion and should not solely be used for evaluation.

Storyline is the least influential out of the three but impactful nonetheless. Because interaction with the storyline element is (compared to interactivity elements) much more passive and the element targets specific (and different) dimensions of immersion and flow, it is a separate element from interactivity. Storyline adds the dimension of enjoyment to immersion and flow, on top of providing for general clarity and comprehension.

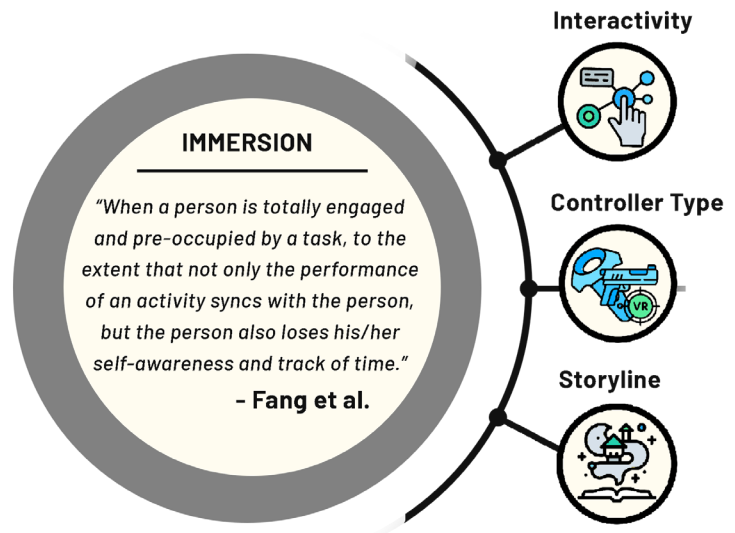


FIGURE 27. IMMERSION AND GAME ELEMENTS

Flow and game elements

Outside of immersion, three other game elements have been identified that targets the dimensions of flow. These elements may help elevate the experience of immersion, lifting it to the experience of flow.

Points, badges and leaderboards adds clarity and enjoyment to an experience. It can help balance challenge and skill and provide for feedback. These elements share the second place with controller type in terms of number of dimensions of flow they tackle. Social interaction contributes to the same dimensions as PBL aside of clarity: enjoyment, immediate feedback and balance of challenge and skill. Levels are the least impactful game element as the dimensions it tackles can also be targeted using other elements. Nonetheless, they are a very widely implemented game element.

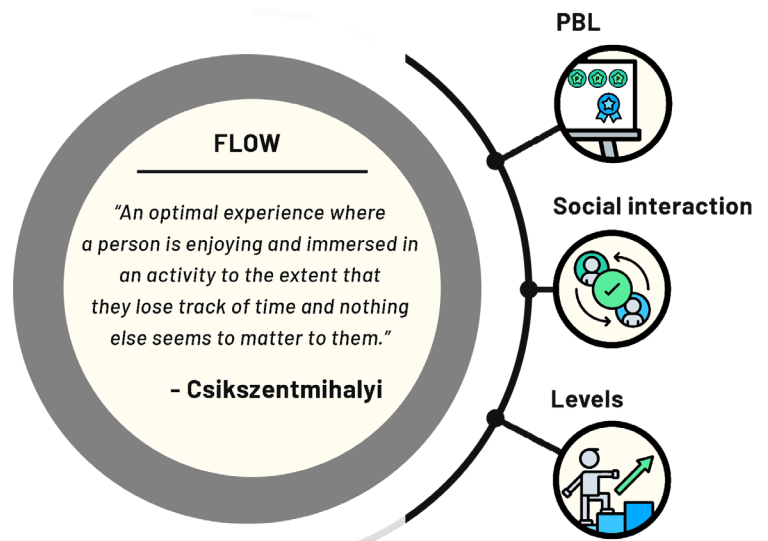


FIGURE 28. FLOW AND GAME ELEMENTS

CONCLUSION

EVALUATION

Because many view immersion differently and the debate on its description is still ongoing, presence has been used to evaluate VR experience in most cases instead, as its description has been established for decades. Although the presence questionnaire can help in evaluation, it does not tell the complete picture. Because symptoms of simulator sickness and usability issues are very common in VR (and very disruptive for overall presence and immersion); in addition to the PQ, researchers often combine the SUS and/or the SSQ to objectively gauge intuitiveness and immersion of an experience. This combination has shown promising results for evaluating and improving VR experiences.

INTERACTIVITY, PRESENCE – IMMERSION – FLOW

The connections between the sub-scales of the presence questionnaire and the elements identified by Radianti et al. show how interactivity is connected to presence. This notion is also supported by the study of Mütterlein (2018), where he concluded interactivity was directly and positively related to both presence and immersion. Because of this, Mütterlein viewed interactivity as a fundamental component of VR. As interactivity is directly and positively related to immersion also, VR elements identified by Radianti et al. have been re-evaluated. Upon closer examination of the sub-scales of the PQ, the possibility to act sub-scale and its basic interaction and assembly elements, and the immediate feedback elements of the self-evaluation sub-scale, fit closely to what is needed for the experience of immersion. Because of this, the relation between the sub-scale and VR elements have been regrouped under interactivity (elements) for immersion (figure 29). Other sub-scales and related elements, specifically the possibility to examine and instruction elements of self-evaluation, have been left as is for presence (figure 30).

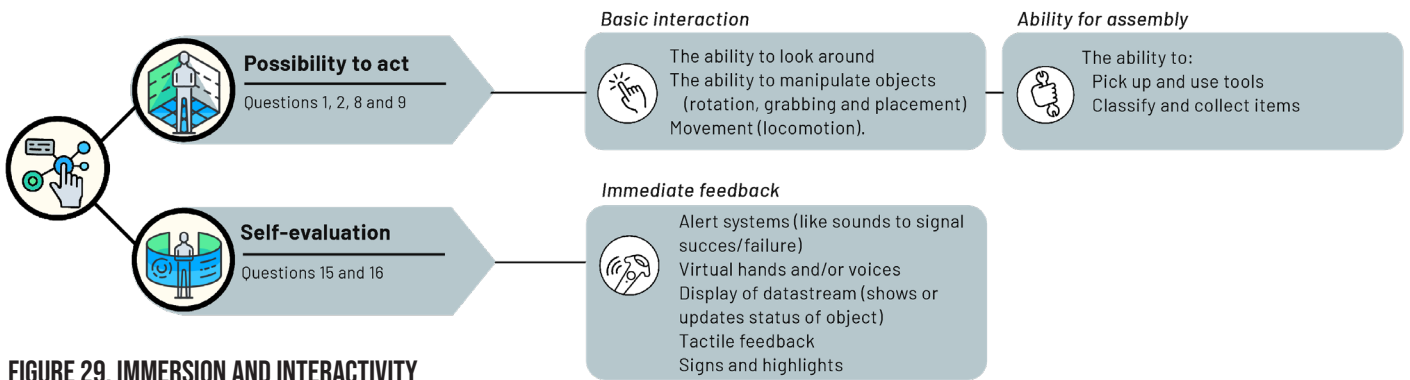


FIGURE 29. IMMERSION AND INTERACTIVITY

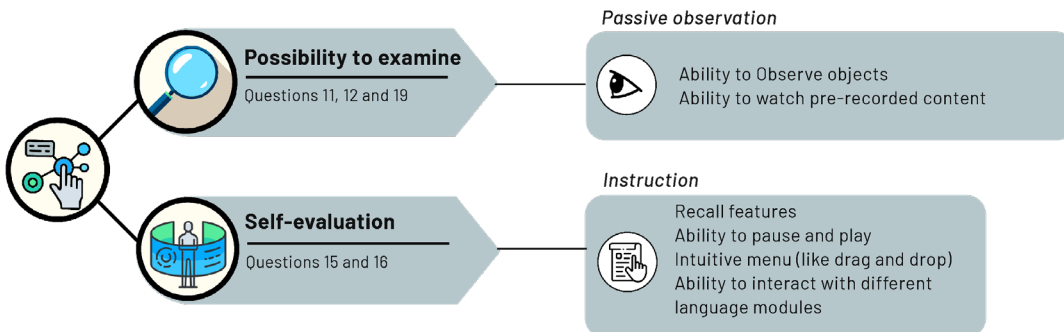


FIGURE 30. PRESENCE AND INTERACTIVITY

There seems to be a degree along the spectrum of how interactivity induces the experience of presence – immersion – flow. Müllerstein's study didn't only show how interactivity is a fundamental component (influencing both presence and immersion, as earlier stated), he also found presence to have a direct and positive relation with immersion. Furthermore, when viewing immersion as a dimension of flow (like Mütterlein did), it becomes evident that immersion itself is directly and positively related to flow, much like how presence is related to immersion. This relation, together with figures 29 and 30, the interactivity elements from section one and the game elements from section two; has been summarized into a visual.

As interactivity is the fundamental component of VR, it can be found as an influencer of both presence and immersion in figure 31. Furthermore, both the technical and psychological point of view of immersion are incorporated into the figure, as the scientific debate on the exact description of the phenomenon is still ongoing.

Elements identified by Radianti et al. are, as previously explained, divided and (re)grouped under interactivity for presence (figure 30) and immersion (figure 29). Other elements impacting presence are derived from the (sub-scales of the) PQ as explained earlier (detailed in dark grey). Aside from the elements of interactivity, other (game) elements have been identified influencing immersion by viewing it as a dimension of flow (detailed in light grey in figure 31). Additional game elements (i.e., PBL, social interaction and levels) can further be used to elevate the experience of immersion and induce flow (seen vertically in figure 31).

Realism for presence and/or storyline for immersion seem similar to interactivity, but are distinctively different due to the elements lying outside of the play area of the player (while elements of interactivity exist inside the play area and often go along with an element of self-evaluation). In addition, realism was also a sub-scale of the PQ (together with quality of interfaces, figure 23) and story was also a separately identified game element (together with controller type, figure 26).

The elements converge downwards towards the player as subjective usability needs and experiences of simulator sickness form a bottleneck for the experience of presence, immersion or flow.

Figure 31 is a visual summary of VR and its components, relations and intricacies. It takes into account established and discussed meanings of interactivity, presence and immersion and also provides concrete examples of how different elements affect the degree of presence and immersion. Using this, the visual can help designers assess VR experiences, but also help gain understanding of the topic VR in general.

Several general take-aways are formulated alongside the visual as general guidelines per level of immersion (presence – immersion – flow). Using these, criteria for intuitive and immersive are proposed.

For players to experience presence:

- Interactivity should include, at the minimum, elements of examination and instruction (to support usability), but can also incorporate more from other levels.
- Quality of interfaces (e.g., an HMD for VR) describes the technology that plays the experience for the player. This medium should, from a technological point of view, be as optimal as possible; no flicker, lag, positioning error and 'disruptively' low screen resolution etc.
- Photo-realism is not required (depending on the experience), but the VR surroundings should be deemed 'realistic enough'.

Although the usability of interactive elements in the experience (perceived interactivity) hinders the experience of presence much less compared to immersion and flow (due to the added complexity of the controller type), simulator sickness is (by contrast) nonetheless a constant potential inhibitor of overall usability. Generally:

- Navigation speeds should be kept relatively low and (sudden) turns and stops should be minimized.
- An 'ease-in' program could help the player adjust to VR in general.

For players to experience immersion, as demonstrated by Müllerstein;

- Some degree of presence is desired (if not required).

Furthermore:

- As possibilities to act and assemble are added for immersion, elements of immediate feedback become more essential to guide and support perceived interactivity
- Controller type is the (extended) medium between the player and the game and should allow for intuitive use; connection and interface issues should be minimized and (tactile) cues should not be misleading.
- Storylines adds clarity and context to the experience and supports better comprehension of general cues and overall enjoyment.

Perceived interactivity and general usability (of the controller type) are bottlenecks to the experience of immersion; as an object needs to be recognized as usable, and the player needs to be able to use the object, before a degree of immersion can be reached.

Therefore:

- An ease-in program to help the player adjust to VR (to help against simulator sickness) and adapt to its interactivity and controls is encouraged.

As immersion is reached, a deeper level of engagement can be achieved by providing (game) elements that support the experience of flow. To allow this, the earlier mentioned usability bottleneck needs to be overcome. Therefore, in order to achieve flow:

- Some degree of immersion is required.
- An ease-in program to help the player adjust to VR (to help against simulator sickness) and adapt to its interactivity and controls is strongly encouraged.

RQ: "What are criteria for making intuitive ("usable") and immersive ("task-engaging") Virtual Reality experiences?"

1. The VR experience does not induce symptoms of simulator sickness, in particular nausea, during gameplay.
2. The quality of interfaces is optimally configured and the VR surroundings are 'realistic' enough, as the player needs to be present before immersion can be achieved.
3. Interactivity elements (specifically for possibilities to act and assemble) have a corresponding perceived interactivity component (i.e., for self-evaluation; like immediate feedback and instruction) as a base for intuitive and immersive interactivity to take place.
4. Controller types are optimally configured and players are able to manipulate interactive elements in the game using the controller (are able to act) to achieve a level of immersion.

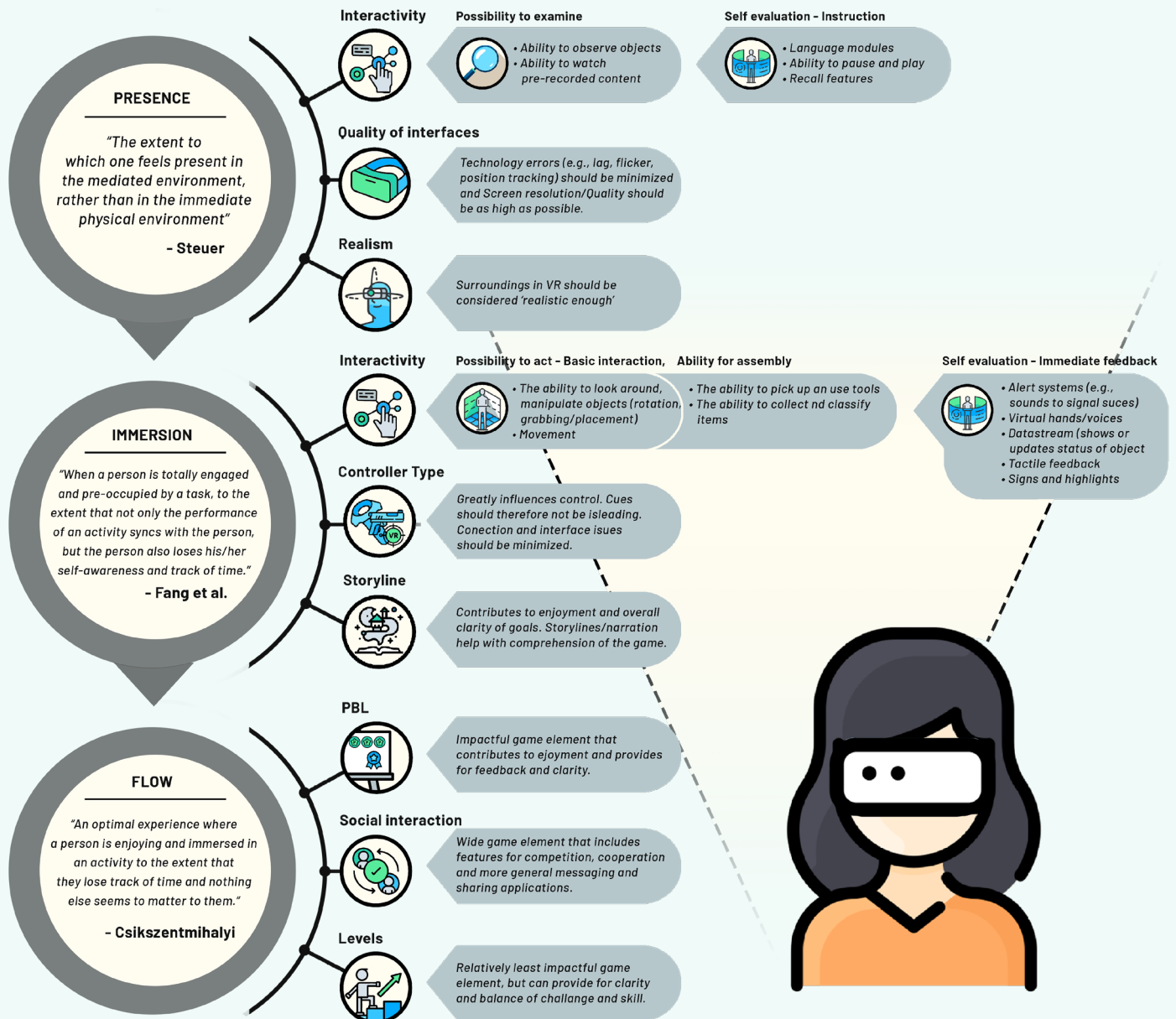


FIGURE 31. LITERATURE SUMMARY; PRESENCE - IMMERSION - FLOW



LIMITATION

Because this literature study was conducted as part of a design project, the time allocated to the study was limited. In addition, SCOPUS was the main database used during collection of relevant papers. The former two heavily influenced the sample of papers analyzed for this study. Also, as the debate on the description and meaning of immersion and the aspects of VR is ongoing, no one-size-fits-all model or criteria can be established. Although the study takes into account several perspectives of immersion, the study mainly hones in on specific aspects of VR; presence, interactivity and immersion.

Synthesis


IMMERSION CRITERIA AND ELEMENTS (& THE VR GAME)

IMMERSION CRITERIA

1. *The VR experience does not induce symptoms of simulator sickness, in particular nausea, during gameplay.* 
2. *The quality of interfaces is optimally configured and the VR surroundings are 'realistic' enough, as the player needs to be present before immersion can be achieved.* 

The VR game does not support locomotion through teleportation and also doesn't manipulate the view of the player. Therefore, if conflicts between the visual and vestibular system are the main cause for simulator sickness, when no interface and controller type errors occur (criteria 2; e.g., low quality, lag, connection issues etc.), symptoms of simulator sickness are not likely to appear. Participants of the preliminary game also did not state to experience any of the oculomotor-, nausea- and/or disorientation-related symptoms. Notably, some participants have stated the game to lag occasionally, which; if not increases symptoms of simulator sickness, definitely decreases the experience of overall presence or immersion.

The quality of the interface used is, certainly in terms of immersion (at least according to researchers like Radianti et al.), a restriction to the level of presence – immersion - flow experienced. The HTC Vive allows for a great quality of interface (in general) when no technology errors occur. The laptop supporting the system is furthermore able to run the VR game exclusively (i.e., run VR with little to no impactful programs in parallel). Notably, although participant didn't specifically state the surroundings to be 'unrealistic' (they were also not directly asked), participants generally mention the game to leave a somber and austere impression. Because of this, this criterion is somewhat, but not fully, satisfied.

3. *Interactivity elements (specifically for possibilities to act and assemble) have a corresponding perceived interactivity component (i.e., for self-evaluation; like immediate feedback and instruction) as a base for intuitive and immersive interactivity to take place.* 

Interactivity is fundamental for VR experiences. Criterion 3 states that for VR experiences to be intuitive and immersive; every element that requires interactivity, also has a component that contributes to its perceived interactivity.

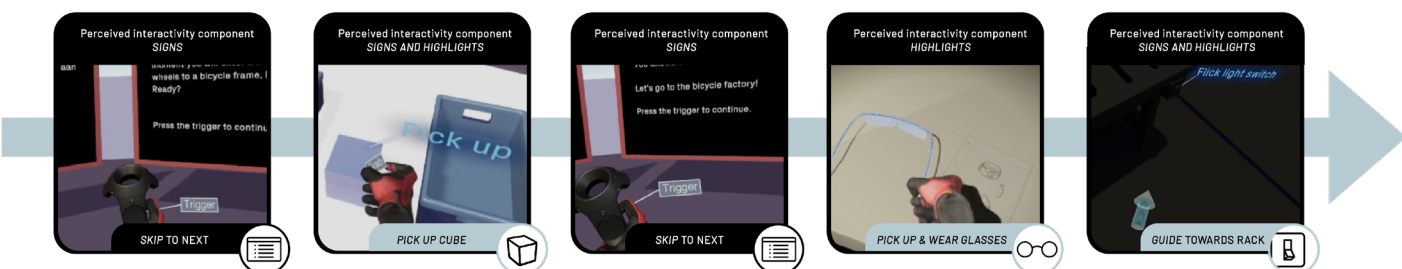


FIGURE 32. INTERACTIVITY ELEMENTS AND THEIR PERCEIVED INTERACTIVITY COMPONENTS IN THE PREGAME

In the tutorial phase of the pregame (figure 32), the user is eased in using two interactivity elements: the board (conveying the main message) and the cube. At the start, when the sequence of the game is initiated, the controller will vibrate in order to draw the attention of the player to the sign (attached to the trigger that needs to be pressed). After, when the cube needs to be picked up; signs (“pick up”) and highlights (blue glow) are used to increase perceived interactivity.

When the player enters the factory, the focus is put on the AR-glasses (literally) by shining a light on it. The glasses are furthermore highlighted upon touch (like the cube in the tutorial phase). When worn, An AR-arrow (feature is called ‘virtual hands’ in figure 31) will appear and guide the player in further tasks (in the main sequence). The arrows are dynamic and, together with the attached sign and highlights, able to point towards the object of interest in much detail (when noticed). After the player flicks the light switch, the main sequence will start.

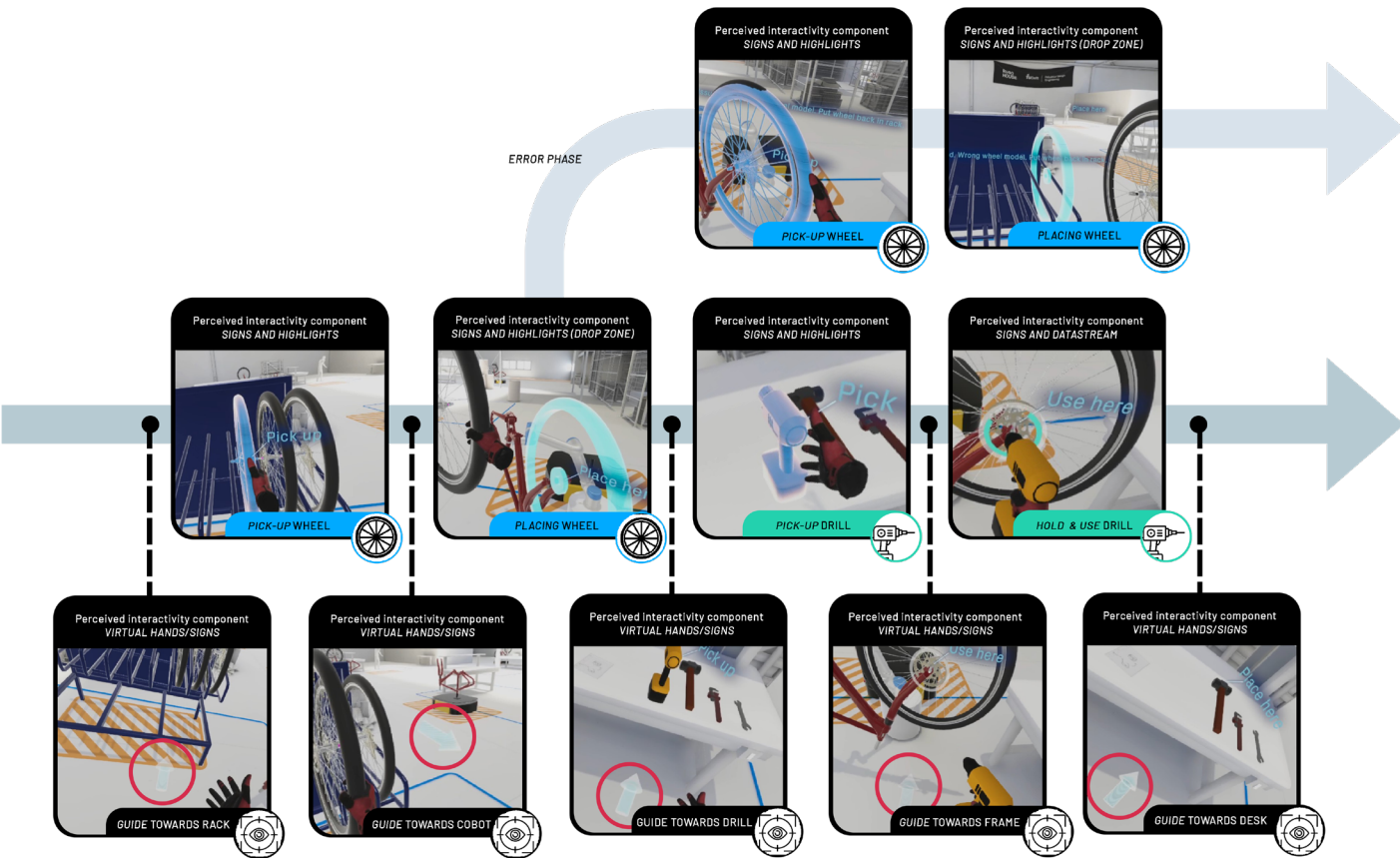


FIGURE 33. INTERACTIVITY ELEMENTS AND THEIR PERCEIVED INTERACTIVITY COMPONENTS IN THE MAIN GAME

For the most part, two elements require interactivity in the main sequence of the game. The wheel has to be picked up and placed elsewhere (blue in figure 33) and the drill has to be used for assembly (green in figure 33). For the wheel and the drill, visual cues are the only components used. The only exception to this is when the drill is activated (green “USE DRILL” in figure 33), where the controller will also vibrate in addition to the sign and data stream shown.

For the wheel (blue in figure 33), signs and highlights are used to increase perceived interactivity. The required action (with the wheel) is communicated through a general message, supported by a sign attached to the wheel; informing players to pick it up or place it. Additionally, high-lights appear to guide the player whenever necessary. For instance, the wheel itself will light up when interaction is required and the object is touched by the player, signaling interactivity. Drop zones will also appear whenever the player needs to place the wheel (and is grabbed) to guide the players.

For the drill, signs and highlights are used in the same manner as with the wheel. Additionally, as the drill is to be used actively (while the wheel is only moved), a data stream showing the status of the action (assembly) is shown to the player. As mentioned earlier, the controller will also start to vibrate at this point.

The VR game, in particular the main sequence; after the AR layer has been activated, incorporates various (mostly visual) components to increase the perceived interactivity of the interactive elements. Although in the tutorial phase of the pregame the elements can benefit from more (focused) engagement, the interactivity elements themselves do take into account how they are perceived as interactive.

4. *Controller types are optimally configured and players are able to manipulate interactive elements in the game using the controller (are able to act) to achieve a level of immersion.*



Criterion 4 represents the bottleneck usability is for immersion (usability of both the controller type and the interactivity elements). As immersion describes the degree to which a player is (optimally) engaged with a task, disruptive controller types (e.g., connection errors, misleading cues etc.) and the inability to act in general will prevent a player to get immersed. Although the earlier criteria are met (somewhat) satisfyingly, and thus the VR game should reasonably be able to at least induce some degree of presence; the same cannot be concluded for criterion 4. Because this criterion is not met, but the previous criteria are, the VR game is likely able to induce some degree of presence, but not immersion.

Although the building blocks for an immersive and flow-inducing VR experience are present in the game; no symptoms of simulator sickness (criterion 1), great quality of interfaces (criterion 2) and interactive elements are present and consider the perceived interactivity component (criterion 3), not satisfying criterion 4 prevents the player from being immersed. This can be observed in the preliminary test for instance; as some players evidently still keep the physical environment in mind while being in virtual reality (error number 9 in the preliminary study). In order to meet criterion 4, the (existing) tutorial of the game should be redeveloped to ease players into the game more comfortably. This should be achieved by focusing on the HTC Vive controller and the players ability to act with it (i.e., basic interaction controls).

IMMERSION ELEMENTS AND THE VR GAME

The lack of control is further substantiated through the usability errors observed in the preliminary test (table 1) and remarks made during the interview (figure 18). Additionally, the game generally leaves an austere (or somber) impression. Players furthermore expressed the need for more clarity during the interview. Figure 31 in the literature study includes elements that addresses different aspects of presence, immersion and flow by contributing in a specific manner. For instance, the storyline element contributes to overall clarity of goals and can help with the general comprehension of the game. Figure 18 compiles the interview take-aways from the preliminary study into one figure. Relevant elements from figure 31 have been linked to these interview take-aways to create figure 34. The lack of control was omitted in this figure because the need for the issue to be fixed was already substantiated.

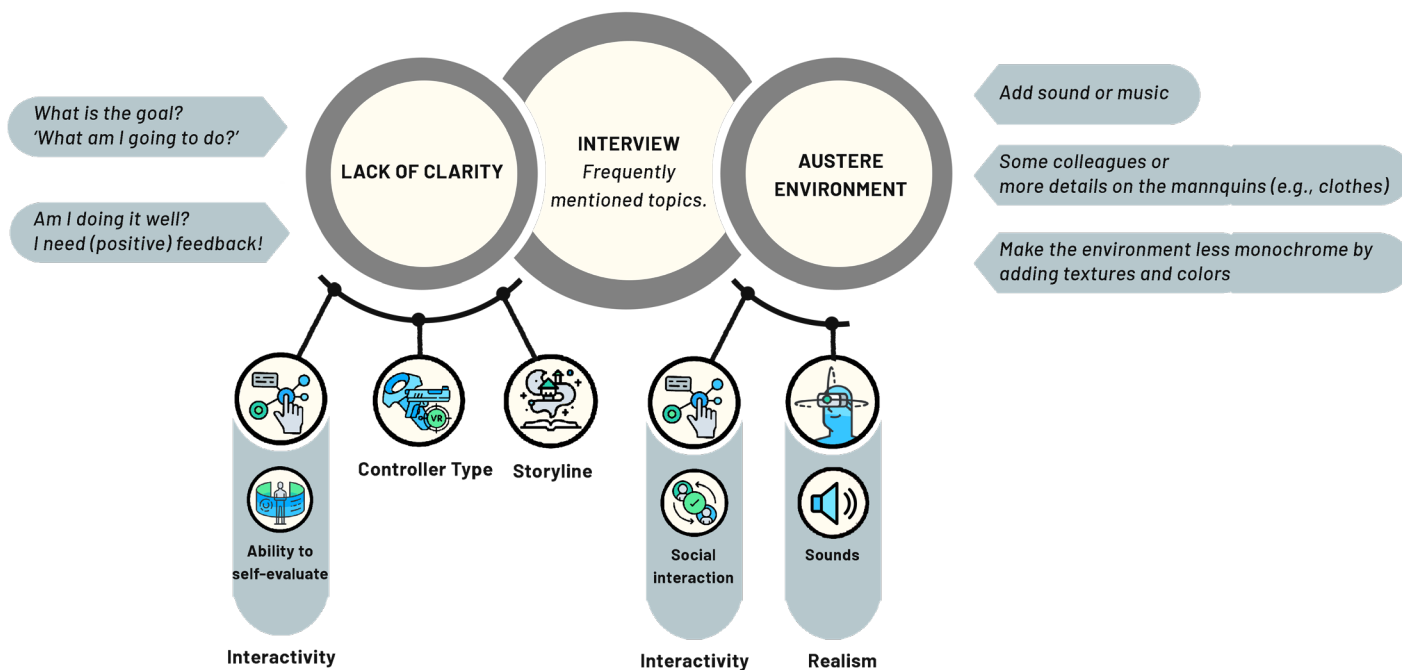


FIGURE 34. PRELIMINARY INTERVIEW TAKE-AWAYS COMBINED WITH RELEVANT LITERATURE TAKE-AWAYS

The austere impression can be solved, reasoned from figure 31 of the literature study; from the bottom up, using realism elements from the level of presence (e.g., background noise/music), or the top down, using interactivity (game) elements from the level of flow (e.g., social interaction). Social interaction like multiplayer or more general messaging features, might solve the issue as players will likely get more engaged with the players rather than the environment itself. Realism can furthermore be improved by adding background noise or music to the environment.

The lack of clarity can be solved using elements to improve immersion. Consequently, this will also support satisfying criterion 4 as clarity and control are not unrelated topics. For instance, usability issues like number 3 observed (during the preliminary test), failing to notice and wear the AR glasses during the factory phase of the pregame, can likely be solved by emphasizing the storyline of the game; using AR in order to collaborate closely with various robots (in the plausible future of work scenario). The storyline element can help clarify the interactivity (of an object) and help players notice cues (solving error number 3). The element adds to the general comprehension of the game (and the story) and is additionally able to improve overall enjoyment. This may also help lift some of the somber or austere impression players get.

During the tutorial phase of the pregame, before the sequence starts and the player is prompted to press the trigger (figure 12 image 1), the controller vibrates to draw attention towards the sign attached to the trigger button (in VR). This cue is confusing for some, as players have expressed not knowing how to respond to it. Misleading controller type cues are generally disruptive and should be minimized; more careful and additional use of self-evaluation elements can also provide for clarity and help players overcome the confusion.

CURRENT VS DESIRED SITUATION

CURRENT SITUATION

Figure 35 is an extended version of the figure found in the preliminary analysis (figure 17). In addition to the total of usability error counted (made by the six participants of the preliminary test), figure 35 also incorporates the relevant elements for interactivity and storyline.

Most players were able to overcome basic interaction errors during or after the second round of assembly in the main gameplay during the preliminary test. Although this happens too late, as the second round of assembly happens in the latter half of the game, it implies that basic interaction errors are easily solvable with some focus.

Additionally, usability errors occur at the start of the game and lead to further issues down the line because the errors are left unresolved. Furthermore, because players struggle during most of the gameplay, the experience is generally regarded as confusing. Also, although the pregame incorporates a message to convey the content (Future of Work), it is debatable whether this is done effectively as many (relevant) cues, like picking up and using the AR glasses to receive extra information, are overlooked.

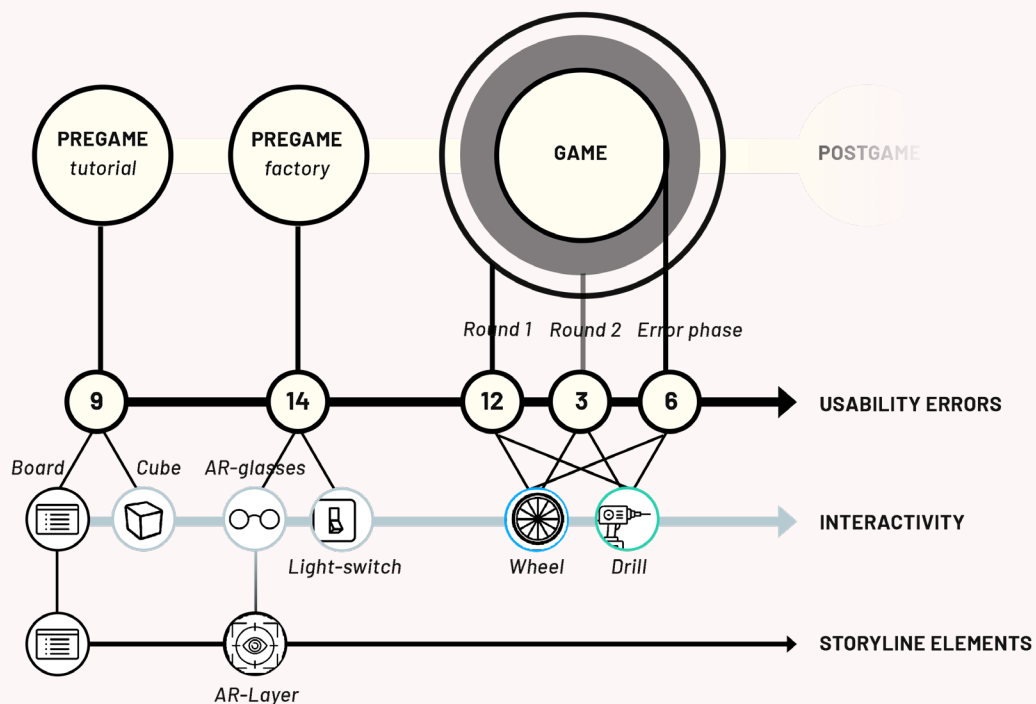


FIGURE 35. CURRENT SITUATION

DESIRED SITUATION

Ideally, the tutorial eases the player in more softly by effectively conveying the storyline and dealing with basic interaction controls. Compared to the current situation, a similar number of errors might occur during the tutorial phase, but once entering the factory, the clarity and usability issues (observed) would drop to zero (figure 36).

Realism of the surroundings is improved by adding interactive background noise but also using sound as a perceived interactivity component. For instance, when interacting with the drill (i.e., holding, using and placing it), various sounds can be used to signal different states of interactivity and provide for immediate feedback. The game furthermore supports having two (or more) players in the same scene.

All proposed improvements are notable, as they are derived from preliminary testing and/or a synthesis of analysis results with literature conclusions. Not all are equally impactful however. Due to the restrictions of this project, a selection is made from the proposed improvements in order to satisfy the time-constraints of this project.

As criterion 4 is the bottleneck for immersion, fixing usability errors will be most impactful. Furthermore, effectively conveying the storyline is part of the *raison d'être* of the game. It also contributes to usability as storylines are able to provide for clarity; which is, as earlier stated, related to (good) usability. Additionally, storylines are able to add to the overall enjoyment of the game, possibly providing a counter for the general austere impression.

Although adding various implementations of sound can contribute to realism and interactivity, it is relatively less impactful than the former mentioned improvements as criterion 3 (regarding realism) is already met somewhat satisfyingly. Furthermore, adding features supporting social interaction will, although likely make the impression less austere, not contribute much to the core issue of preventing players from getting immersed. Because of this, improvements regarding the VR game will focus on two 'jobs-to-be-done' (JTBD) in the tutorial phase, specifically; conveying the story of the plausible future of work scenario and teaching basic interaction controls.

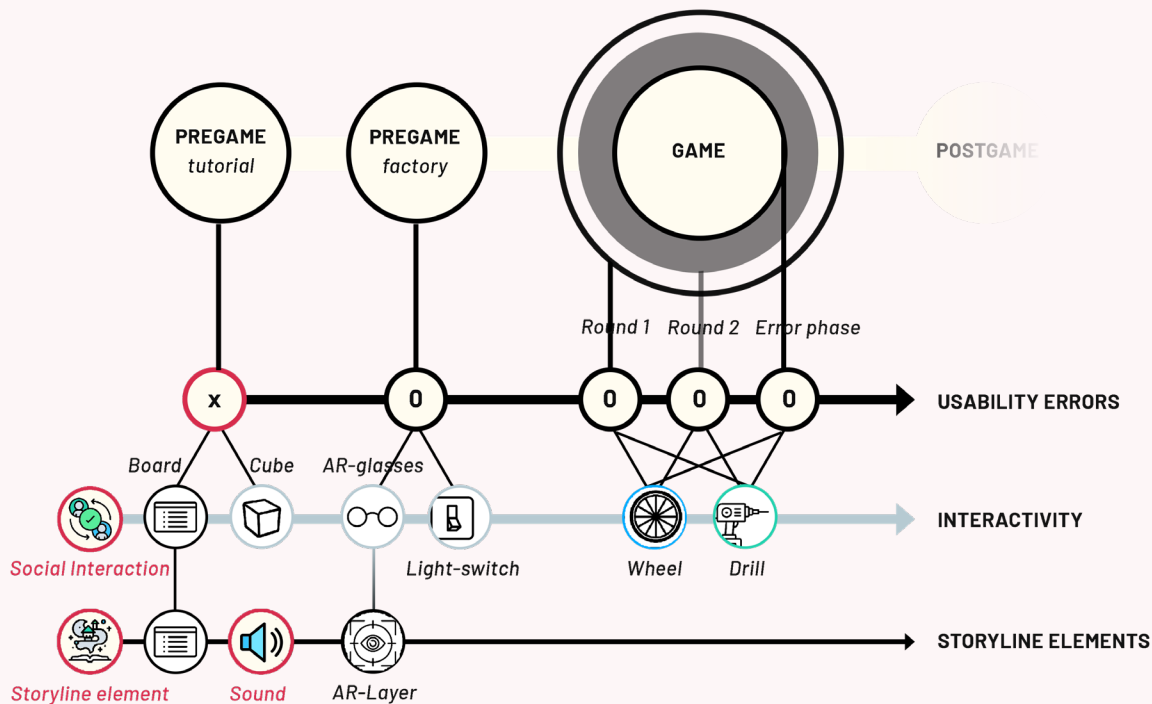


FIGURE 36. DESIRED SITUATION

HUMAN CENTERED APPROACH

In addition to immersion being a subjective experience, the JTBD; conveying the story and teaching basic interaction controls in VR effectively, are also very subjective and dependent on user characteristics. Different users comprehend and learn very differently because of variation in abilities, needs and skills. A human centered approach is therefore preferred in order to resolve the JTBD satisfyingly. According to ISO (ISO 9241-210, 2019); the main principle of human-centered design involves the inclusion of *relevant users* in development, and refine/iterate on solutions and fixes according to user *evaluations*.

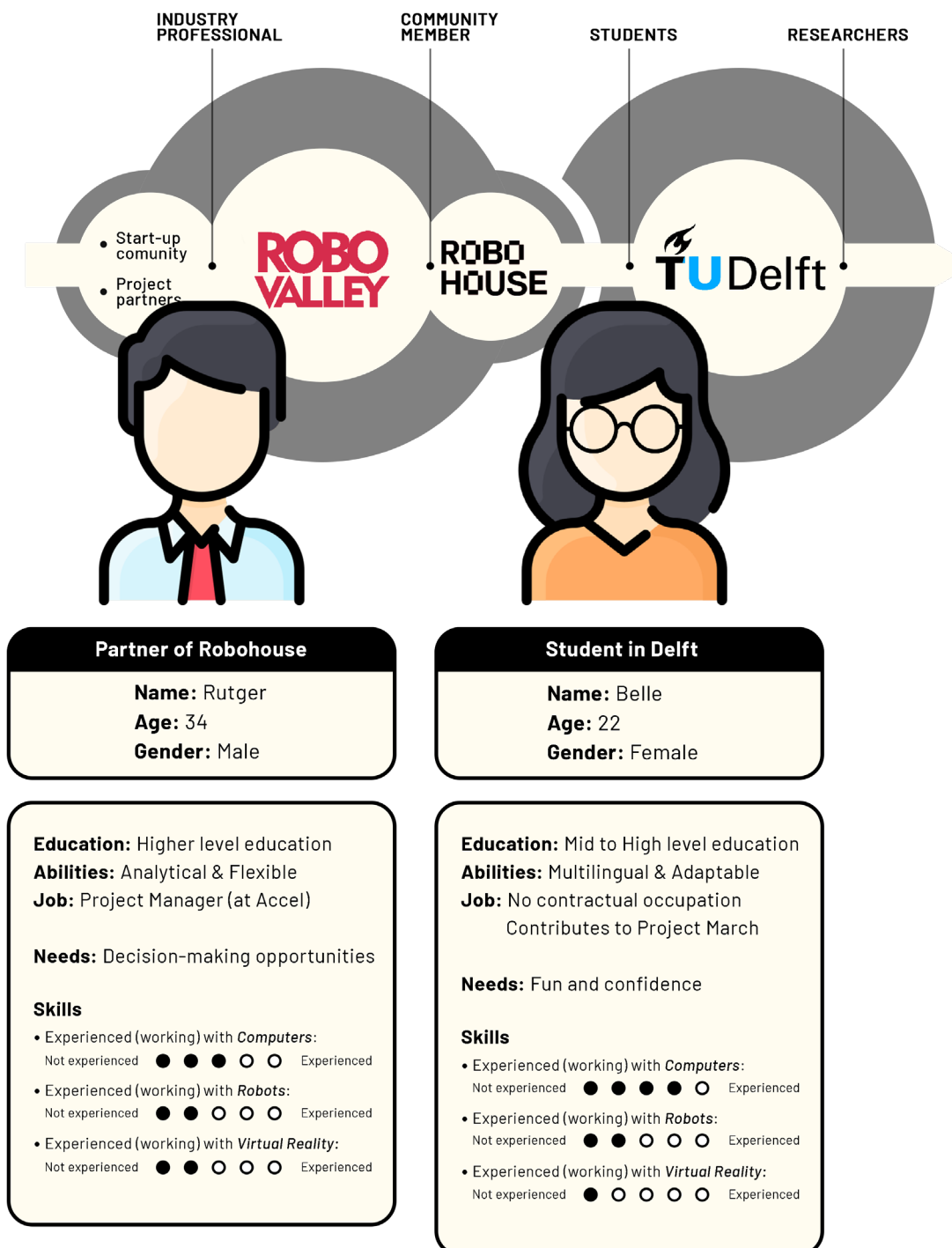


FIGURE 37. PERSONA, SYNTHESIZED FROM RELEVANT USERS AND (EXTRA) INTERVIEWS

PERSONA

Relevant users were identified earlier based on the context of use (figure 11). Two personas have been made based on this together with a (re)cluster of the interview taken in the preliminary analysis (appendix B) and additional interviews conducted with industry professionals (i.e., workers at a bicycle company). The contacts funnel of the industry professionals approached, along with the interview setup and minutes, can be found in appendices C and A respectively.

The personas give a rough estimation of the different relevant users, although generally they are much alike. Because of this, community member and researcher has been left out as the former (also) consists out of students and/or industry professionals and the latter is similar to students and (industry) professionals.

EVALUATION

Three evaluation methods stand out when it comes to making refinements and improvements in VR (as they have shown promising results); using the presence questionnaire to evaluate presence according to its sub scales (figure 23 can help suggest for improvements when doing so), using the system usability scale to score usability in a 'quick and dirty' manner, using the simulator sickness questionnaire and its sub scores to gauge the degree of simulator sickness experienced.

As criterion 1 has been met satisfyingly, the use of the SSQ is not necessary.

Interactivity, in particular, the possibility to act, is included in the sub scales of the PQ. Because of this, the questionnaire is able to (also) evaluate immersive experiences; as the possibility to act sub score will be higher for an immersive experience compared to an experience where players only feel present.

Criterion 4 revolves around usability being the bottleneck for immersion. Currently this criterion is not met and the main reason why players experience troubles getting immersed. Although using the PQ to validate presence/immersion will give more insight, as the results involve five sub scale scores; due to the game meeting all criteria except for criterion 4, using the SUS allows for a 'quick and dirty' assessment into the degree of usability being a bottleneck for immersion in this game.

Developing the experience alongside the relevant (end) user is more favorable however, as the experience can tailor itself to meet the highly subjective needs of the JTBD. The Rapid Iterative Prototyping and Evaluation (RITE; Medlock et al., 2007) method is a usability testing framework that conforms to the ISO norm (2019). The method incorporates relevant user tests and evaluations to iterate on usability fixes.

In order to meet criterion 4 and allow players to get immersed into the content of the VR game, the tutorial phase of the pregame will ease players in more easily by emphasizing the story and the basic interaction controls in the game. Because of the subjectivity of the JTBD, fixes will be implemented and evaluated according to the RITE method. Additionally, the SUS score can provide further insights into usability.

What stands out is that participants had the need to compare in order to 'get the job done'. Specifically, explaining current and future scenarios would clarify the differences of working (in the future and now) and help set the context for the story. Furthermore, 'seeing' someone else perform and mirroring the behavior was a commonality in many of the ideas created (e.g., dedicated YouTube channel for tutorials, film VR experience prior, have a virtual assistant etc.).

It seems that using an element that supports this; for instance, guiding players through watching various video-examples, would be sufficient to fulfill the JTBD and consequently criterion 4 (so players can get immersed). However, further testing and development is needed to validate (and iterate) this.

THE RITE METHOD

The RITE method is a usability testing framework that was developed by Medlock et al. (2017). The method was implemented (and validated) during the development of the game Age of Empires II. The main principle of this method revolves around the user evaluation of (the efficacy) of a fix, immediately after implementation. It was developed in order to meet the goal "to ship an improved UI as quickly and cheaply as possible".

Traditionally, user evaluations are used to uncover (preferably all) usability issues. Fixes can consequently be proposed and (hopefully) implemented in order to solve the issue. However, fixes often do not get implemented and studies have shown several reasons for this. For instance, teams are unsure whether the fix will solve the problem or when usability feedback arrives relatively late, the delay might act as a barrier for the implementation of the fix etc. (Medlock et al., 2017).

When using the RITE method, the focus is laid more heavily on whether a problem is fixed and to determine what the likelihood is that the fix solves the problem, rather than "agonizing over whether every problem has been uncovered" (Medlock et al., 2017). Because the main issue (not satisfying criterion 4) already has been uncovered and the JTBD are so subjective; the RITE method can help develop and iterate on fixes in order to solve the problem effectively.

RULES

Similar to 'traditional' usability testing methods; a target group for testing and what and how issues will be measured needs to be defined. Furthermore, a test script and suiting prototype needs to be constructed. The RITE method differs from traditional usability testing due to the emphasis laid on evaluating the likelihood of a fix solving an issue. In order to achieve this, fixes to issues are proposed and implemented according to the level the issue is classified in. There are four categories:

1. "Issues that have an obvious cause and an obvious solution that can be implemented quickly.
E.g., text changes; re-labelling buttons, rewording dialog boxes etc.
→ Implement the fix and use revised prototype with the new participant
2. Issues that have an obvious cause and an obvious solution but cannot be implemented quickly/within the timeframe of this test.
E.g., (new) features that require substantial design and or code changes etc.
→ Start implementing the fix and implement the revised prototype as soon as it is finished.
3. Issues that appear to have no obvious cause and therefore no obvious solution.
4. Issues that may be due to other factors
E.g., test setup, interaction with participant, etc.
→ Collect more data in order to upgrade to either category 1 or 2, or classify as 'non-issue.'" (Medlock et al., 2017)

TEST SETUP

The goal is to implement fixes to the tutorial part of the pregame in order for players to grasp the basic interaction controls (before entering the factory) and have an overall more cognisant experience of the content of the game, easing them into the experience more comfortably. In order to support this, three types of data will be gathered: Observation (1), Interview (2) and a System Usability Scale score (3).

1. Observations from the researcher will be used to evaluate usability, specifically:
 - a. Errors: Player is confused/doesn't grasp the task at hand, but somehow manages to proceed.
 - b. Failures: Player can't advance on their own (and needs instructions from the researcher; added as clarification after the pilot).
2. The interview will be used to evaluate the player's grasp on the content of the game, but is also an opportunity to inquire more about errors, failures or other observations (if necessary), as players are encouraged to 'think out loud' during gameplay.
3. Because fulfilling the JTBD will, although likely, not certainly equate to better usability, the SUS Score might give more insight into this topic.

Eight usability errors have been identified during the preliminary analysis (table 1), of which the first two occurs during the tutorial phase. These two have a zero tolerance for failure as, due to the linearity of the game, failing them would prevent the user from proceeding into the game (and cause usability issues later on). Additionally, the third error also needs to be solved without exception, as using the AR-glasses (and the AR-layer it creates) is part of the story the game conveys.

1. Using the controller to press the trigger in order to advance to the next scene
2. Using the controller to press the button in order to grab (or release) an object (block)
3. Using the controller to press the button to grab (and hold) the glasses to put on

The other five errors observed are symptoms of the lack of control (and clarity to a lesser extent). These also have a zero tolerance for failure, but will be dealt with indirectly in the tutorial phase.

4. Using the controller to press the button to grab (and hold) the drill.
5. Using the controller to press the button to activate the drill (while grabbing)
6. Interacting with a game object (grab and rotate the priorly attached wheel) in VR
7. Interacting with a game object (grab the priorly attached wheel) to put somewhere else
8. Identifying the object (grab last wheel on newly ordered rack) that needs interaction in VR

STIMULI

Qualtrics questionnaire

Qualtrics will be used for the SUS questionnaire. Additionally, some general information about the participant will be gathered (name, age, occupation, education and skills as described in persona), along with the participant's consent (to use collected data of the test). Screenshots of the questionnaire can be found in appendix F. The general flow can be seen in figure 39

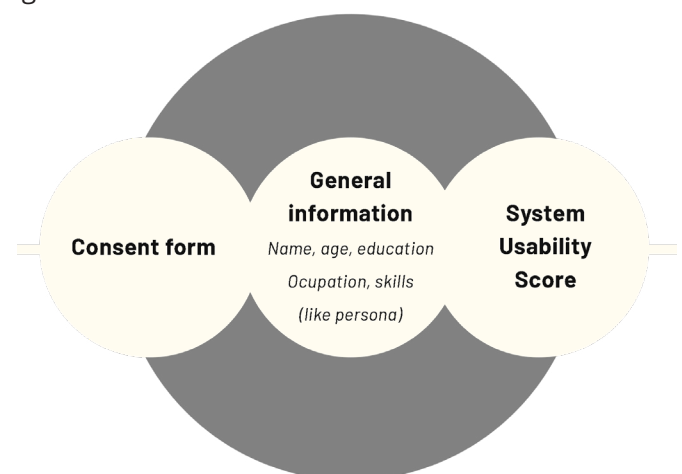


FIGURE 39. QUALTRICS FLOW

Interview script

A script will be used to guide the interview in order to inquire the participants consistently and completely. The data collected will be used to help classify issues observed and revise the prototype. The script can be found in appendix A. Consent and minutes of the interview can be found in appendix H.



FIGURE 40. PILOT PROTOTYPE

VR game prototype

In order to show the 'Future of Work' (FoW) storyline and explain the controls by example; a mixed reality (MR) clip, including both digital elements (to show the FoW environment) and physical elements (to show someone controlling the environment) combined into one display, will be used. Capturing the gameplay from a third perspective provides support while controls are being explained, as players are able to mirror the video. Additionally, viewing the gameplay from a third perspective gives the player an overall view of the surroundings before entering the space in a first-person view. This allows the storyline to be set in a relatively richer manner.

Using existing elements and text, supported by a MR clip, will likely ease the player into the game, before getting engaged with the tasks (i.e., before getting immersed) in the main sequence, much more effectively.

A table with two tablets is added (stationed before the two existing boards in the tutorial phase) to show two mixed reality clips. The first clip shows a player in the main sequence making an assembly; grabbing and dropping wheels and attaching it using the drill. The second clip shows the player in the factory phase of the pregame. The focus is put on activating the AR layer using the AR glasses in this clip.

Additionally, a table with a drill is added for players to explore if they desire to do so. While the sequence does not require players to interact with the drill, adding the most complex interaction (as the preliminary test results have shown) into the tutorial phase, after showing the MRC examples, might help

players get over the usability hurdle. Usability error number 9, observed during the preliminary test; players tend to not physically move much during gameplay, is likely a result of players not being immersed. Thus, the issue is likely to be solved indirectly (through meeting criterion 4/JTBD). Nevertheless, the initial table players interact with, the table with the four containers and the cube, has been moved to the other side of the play area. After the story and controls are explained, players need to walk to the other side of the play area in order to conclude the tutorial phase.

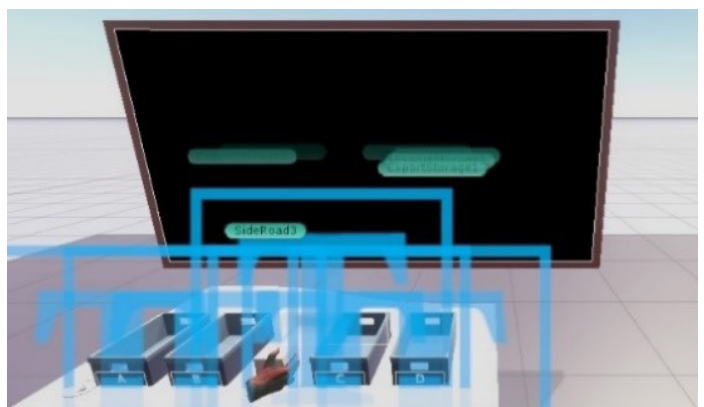


FIGURE 41. SEQUENCE RELOCATION

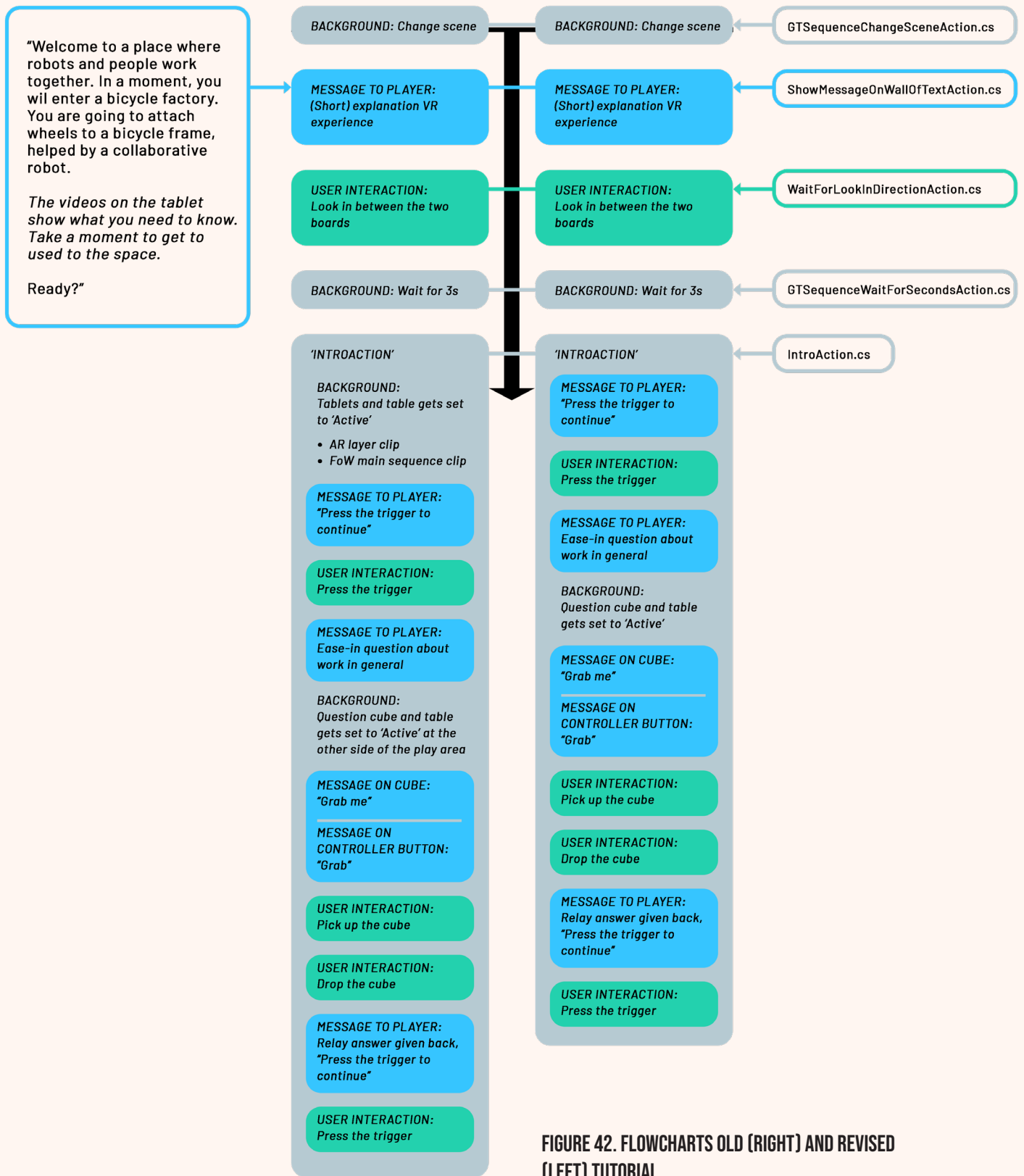


FIGURE 42. FLOWCHARTS OLD (RIGHT) AND REVISED (LEFT) TUTORIAL

Figure 42 is a flowchart of the tutorial phase of the pregame (flowchart of the whole game can be found in appendix D). On the right, the unedited tutorial is displayed along with (parts of) the coinciding C# scripts. On the left, the edited version of the tutorial is can be seen. Here, the IntroAction.cs script has been added to incorporate the added scene-elements into the flow. The messages have mostly been left unchanged and can be viewed (for both versions) on the left side. The (new and) edited scripts, along with all the rest of the media can be found in the 'PakGrad' folder in 'Assets'.

Preferably, as the desired scenario states, basic interaction control issues are not encountered anymore after entering the factory. As only one round of assembly is needed to validate this, the entire game has furthermore been shortened: the edited tutorial phase of the pregame – factory initiation phase – Single assembly (of wheel to frame) in the main sequence; without the error phase.

PROCEDURE

The test setup consists out of two phases. In the first phase, the participant will use the VR prototype and think out loud, while the researcher observes the gameplay (and minds the list of actions with zero tolerance for erros). The researcher will take place behind the laptop (1) and the participant will play within the VR area created by the HTC lighthouses (3a in figure 43). This part of the test will be recorded using a camcorder (4).

- *"You are going to play a game in Virtual Reality for a usability test. In the game, you will be put into a possible scenario of what the future of work might look like. Think of it as playing in a world a couple of decades from now. Please think out loud during gameplay. Are you ready?"*
 - Prepare tester for the experience generically (tell as little as possible when it comes to the content [storyline/narration] and use [usability]). When asked, tell the player that it'll all become clear when the experience starts.

- The researcher helps the tester with putting on the HMD and hands over the controllers. The researcher then starts the game by pressing play in Unity.
 - During this phase, the researcher will step in when necessary. Errors and failures (as defined earlier), will be noted down together with the part of the scene in which it took place. Later, during the interview, more information can be inquired.

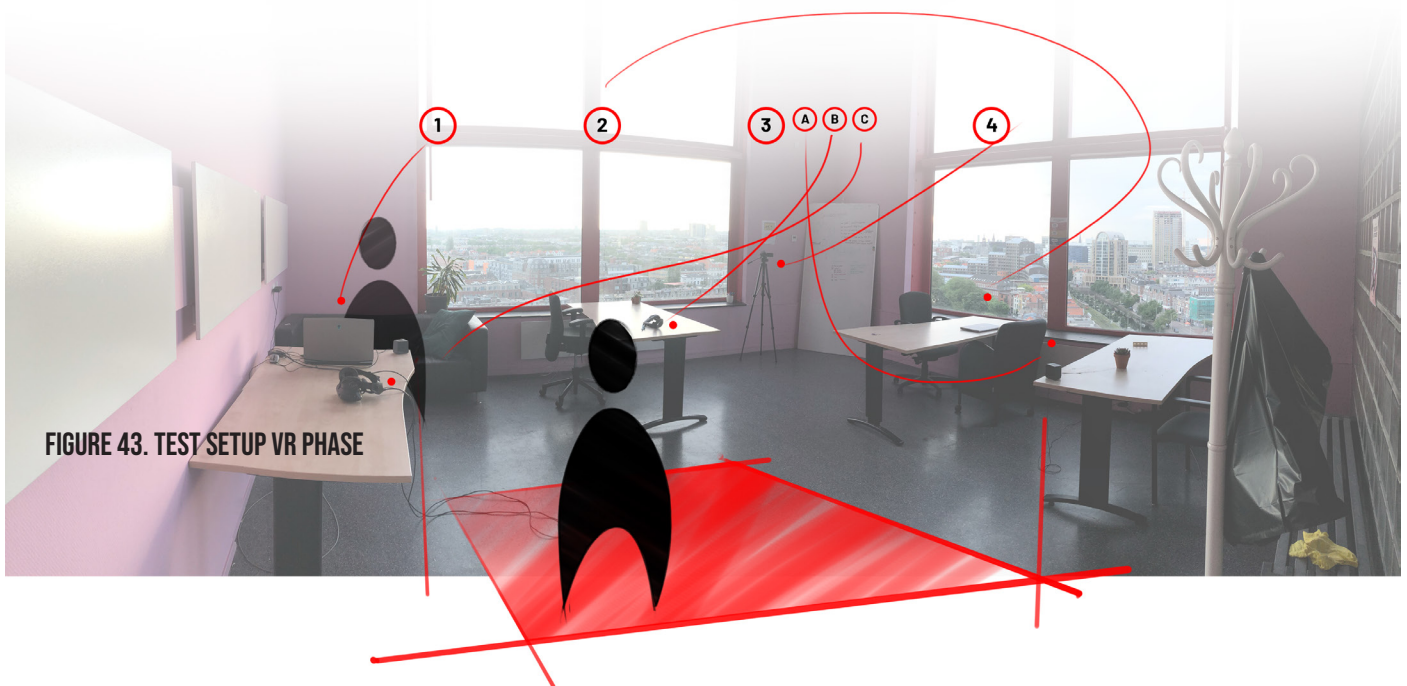


FIGURE 43. TEST SETUP VR PHASE

When gameplay is finished, the second phase of the test start (figure 44). In this phase, the participant will be interviewed by the researcher (following the earlier mentioned script) and rate the overall usability using the SUS questionnaire (on Qualtrics). The researcher helps the participant put away the controllers and take off the HMD and goes towards the second laptop (2), so the participant can be interviewed and fill out the questionnaire after. The entire procedure can be seen in figure 45.

Although a script has been set up (and the question certainly needs to be asked to all participants), the interview should also allow for a natural flowing conversation. Let the participant talk as much as possible. When an answer to a question is vague, the same question will be repeated with some words replaced (with synonyms).

An Alienware laptop (1) will be used to run Unity (version 2019.4.10f) and the VR game prototype. 'OBS Studio' will be used to capture the gameplay (omitted after the pilot). Another laptop (2) will be used, in order to relieve potential (over)load, to take notes (during observations and the interview) and for the player to fill in the SUS questionnaire (in Qualtrics). The HTC Vive setup (version 1, including 3a: two lighthouses and 3b: two controllers along 3c: an HMD) will be used to play the game, which will be recorded on a HD camcorder (4: Panasonic HDC-SD20, 1920x1080 Full HD).

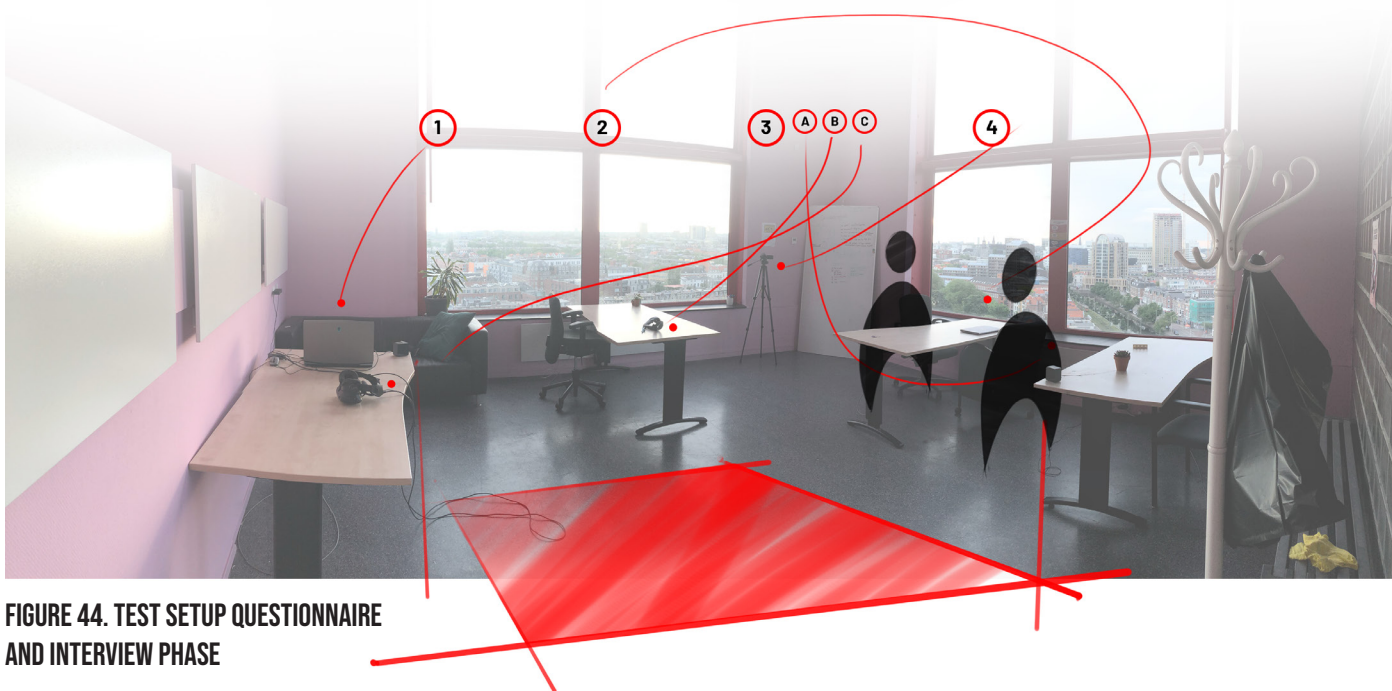


FIGURE 44. TEST SETUP QUESTIONNAIRE AND INTERVIEW PHASE

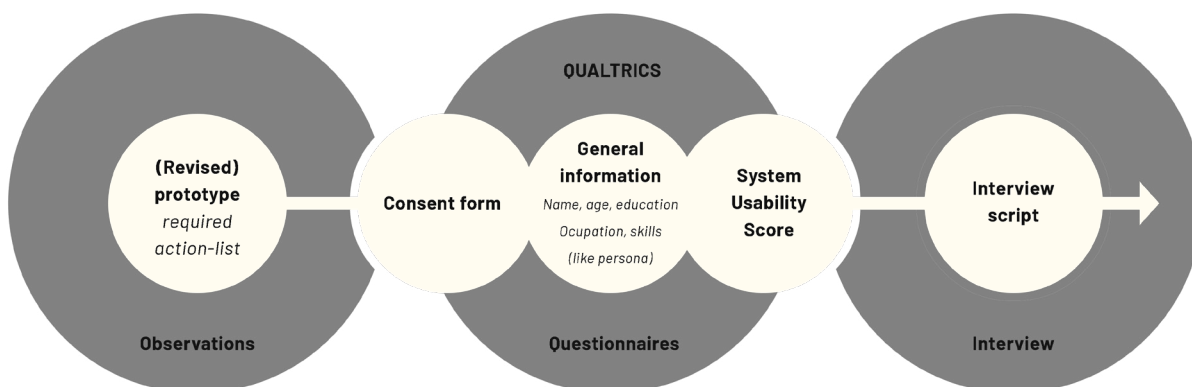


FIGURE 45. TEST FLOW

TARGET GROUP

The users of the game are currently general visitors of Robohouse and students around the faculty of Architecture/on the campus of TU Delft; because of the context of use, as explained earlier. More generally speaking, these types can be summed under "Individuals visiting Robohouse who have little experience with Virtual Reality, but are interested in trying it". As the pandemic makes it impossible to test on site, preferably with individuals of the Robohouse community themselves, they will be simulated using students, which is at least one of the two subsets of the users of this game.

Sample sizing

Generally speaking, the literature (according to Medlock et al., 2007) suggests that running zero participants will uncover zero problems and the more participants are used, the fewer new problems get discovered. Turner et al.'s 'cumulative binomial probability formula' (Turner et al., 2006) is used to determine the right sample size for A/B tests when the likelihood of detection for a problem is known.

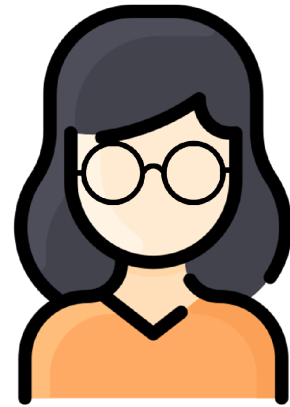
The individual problem discovery rates (P) of the preliminary study ranged from 0.22 to 0.66, averaging on 0.39 for that sample. Because small sample estimates of P are almost always inflated, Turner et al. proposes a method for an accurate assessment of true proportions of problem discovery rates. Following this method, the problem discovery rate averaged on P = 0.22 (appendix I)

Although an exact sample size can be calculated using this average; through general rules of thumb provided by Turner et al., it can already be concluded that at least 5 participants need to be observed, as the various problem discovery rates never exceeded 0.4 (not on average nor through any of the further manipulations). Using 5 participants is furthermore advised by Nielsen (2012)

- When the P is between 0.3 and 0.4, observing 4 – 5 participants will uncover roughly 80% of the problems in a user interface
- When the P is very high (>0.5), observing 1 – 3 participants will uncover 87.55 of the problems in a user interface
- Problems in a UI that, for whatever reason, doesn't have a high likelihood of detection will require more participants.

Following the method, 16 participants were subsequently calculated (for a 99% rate of all problems being detected). Because the method is mostly used to suggest for sample sizes in A/B testing; where the problem discovery rate stays constant, 16 is likely to be a bit much.

Nonetheless, because a sample size of between 5 and 16 participants seems reasonably appropriate, the RITE method aims to incorporate at least ten students. However, more or less might be needed depending on the efficacy of the implemented fixes (and the resulting number of errors and failures observed).



Student in Delft

Name: Belle
Age: 22
Gender: Female

Education: Mid to High level education
Abilities: Multilingual & Adaptable
Job: No contractual occupation
 Contributes to Project March

Needs: Fun and confidence

Skills

- Experienced (working) with **Computers:**
 Not experienced ● ● ● ● ○ Experienced
- Experienced (working) with **Robots:**
 Not experienced ● ● ○ ○ ○ Experienced
- Experienced (working) with **Virtual Reality:**
 Not experienced ● ○ ○ ○ ○ Experienced

FIGURE 46. TEST PERSONA

ITERATIONS

PROTOTYPE 1.0: PILOT (ROUND 0)

A pilot has been conducted with a student following the procedure as described earlier. During this round, the following failures and errors were observed:

Failures observed:

1. Failed to (notice and) use the trigger to advance to the next scene.
2. Did not immediately understood to hold the grip button in order to grab the cube.
3. Did not specifically notice the AR glasses as interactable.

Errors observed:

1. The start of the sequence was unintentionally triggered by the player.
2. The player needed direction to notice the wheel (to be placed).
3. The player needed direction to hold and activate the drill.
4. The player did not understand the difference between VR and AR (during gameplay)

The player was generally disoriented and distracted by all the objects at the start. The sequence was furthermore unintentionally triggered by the player; activating the table with the tablets and the drill. All notable failures (as stated in the test setup) occurred and more. The player needed much help from the researcher to get through the gameplay.

The storyline, collaborating with robots with the help of AR, was not conveyed effectively as neither AR nor collaborating with robots have been mentioned when inquiring about the content. Also, there is a misunderstanding in wearing the HMD to experience the game in VR and wearing the AR glasses (in VR) to activate the AR layer. The player was not able to make a distinction between the two technologies.

Even though the participant encountered many usability errors during gameplay, SUS indicated nonetheless the usability 'to be okay, but could use improvement'; with a score of 65. Notably, the participant mentioned to 'finally be able to try this out', stating her desire to use VR in general. This falls in line with observations from the preliminary results, where participants also generally mentioned to be satisfied with the experience of the technology, even though the content was troublesome to get through.

Issue classification

Clearly, more explanation is needed to teach the basic interaction controls. Also, the sequence gets triggered when the player looks in between the two boards currently. This trigger needs to be improved so players don't unintentionally start the sequence. Because errors 2 and 3 observed are located in the main sequence of the game, solutions are not classified in the table; as the errors should be resolved during the tutorial.

The colored rows are failures that were mentioned in the list of 'actions all participants need to be able to go through, without exception'. These issues have a relatively higher priority compared to others, as they are closely related to the JTBDs and the goal of the redevelopment.

Level 1 Classification		
Round	Description	Issue
Pilot (0)	Failure 1	MRC about controls was not able to help the player locate (and use) the trigger button on the controls.
Pilot (0)	Failure 2	Was not able to intuitively (without external intervention) use the grip button to grab the cube during the tutorial.
Pilot (0)	Error 1	As the user was looking around to get to know the surroundings; the trigger in order to start the tutorial, placed in between the two boards, was unintentionally activated.
Level 2 Classification		
Pilot (0)	Failure 3	MRC clip did not help the player notice (the cue on) the AR glasses.
Level 4 Classification		
Pilot (0)	Error 4	Player was not able to distinguish between AR and VR technologies during the interview.



Improvements

The newly added tables with drills and tablets contributed a lot to the confusion the player experienced. The player was not able to focus her attention as too much was going on at the same time. Furthermore, as the tablets appeared within the play area of the game, it was perceived as interactive while it was not. Evidently, when it comes to clarity of the job to be done; *less is more*.

1. Both tables containing the drill and tablets will be removed.

The MR clips were perceived as vague. The player noted to not like the perspective of the clips (recorded from the back). Additionally, the clips were not able to convey the need of the AR glasses nor the use of the controllers effectively. As the tablets are removed, the board will be used to support the clips. Although the content clearly needs to be reshot, additional messages and sprites will be used in order to help explain the basic interaction controls.

2. The existing boards will be adjusted in order to support the MR clips for the controls and the storyline (figure 49).
3. Sprites will be added alongside more messages to guide the player in using the grip and trigger buttons (figures 50 and 51 respectively).

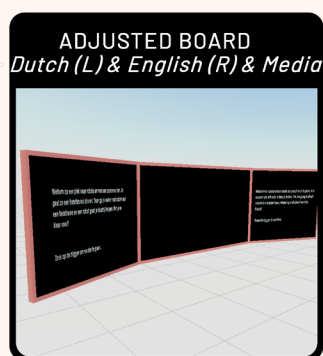


FIGURE 49. BOARD IMPROVEMENT



FIGURE 50. SPRITE FOR GRAB BUTTON



FIGURE 51. SPRITE FOR TRIGGER BUTTON

Lastly:

4. The trigger in between the board will be relocated to the ground so the player needs to 'nod' in order to proceed.

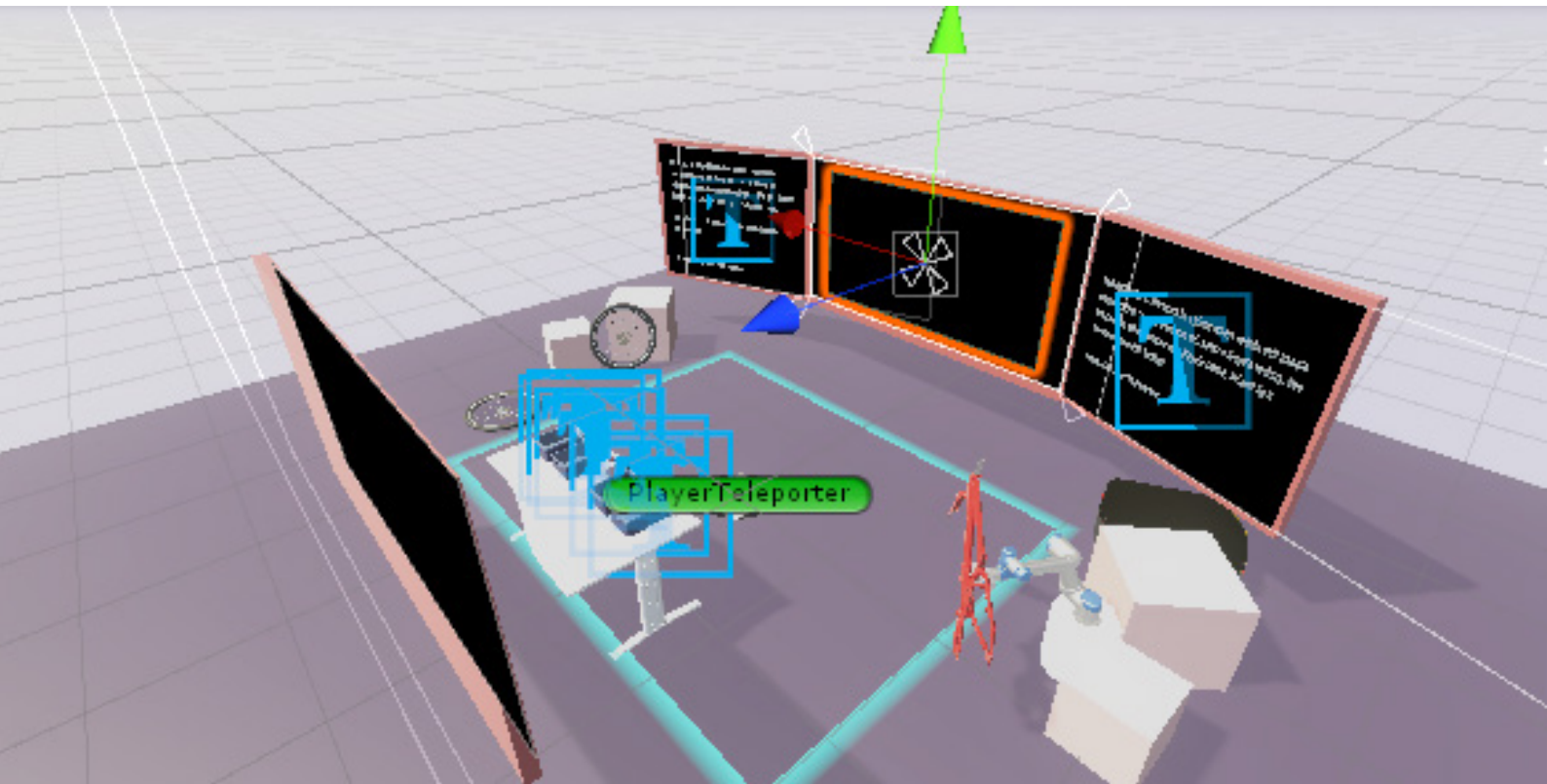


FIGURE 52. PROTOTYPE V1.1

PROTOTYPE 1.1: RITE ROUND 1

During this round, the board has been adjusted to allow for Dutch and English texts, displayed left and right respectively, together with big MR clips and sprites, in the middle. Two more messages are also added to guide the player and provide for positive feedback after the right button has been pressed. The table with the cube and containers, together with the board, are kept at the back of the player (from the pilot round). The scene can be seen in figure 52.

During this round, the following error was observed:

1. Couldn't smoothly grab the cube to answer the question. Did not need external instruction however.

Using the sprites together with the messages explain the basic interaction controls much more effectively compared to the prototype used in the pilot round. Notably, the player did not have issues using the drill during the main sequence, but did have troubles performing a grab (an arguably much less complex interaction) during the tutorial. This is likely because there is relatively more emphasis on the use of the trigger in the tutorial compared to the grab button (pressing the trigger is needed to advance further in the tutorial, pressing the grab button isn't).

Furthermore, the player is overall less confused during gameplay. The "less is more" approach makes the tutorial phase much clearer. In line with this, the gameplay should support one language rather than two, as the latter takes up much of the 'visual real estate' (while not being interacted with). As the messages given during the main sequence of the game are English, the tutorial will support this also. After redevelopment, the entire game (including the pre- and main-game) can be built into different languages.

As the proposed (level 2) reshoots haven't been implemented in this revision yet, similar to the pilot round, the MR clips did not suffice to convey the storyline effectively. However, the player was able to mention one video to show the context after further questioning from the researcher. The correlation with the future of work was eventually made after the player was inquired to project her vision of the current manner bicycles get assembled.

The SUS questionnaire scored 88; 23 points higher compared to the pilot round, passing the limit of 80.3 points (meaning people love your system and would recommend it to their friends) comfortably. It was expected that the score would be relatively higher due to the player not encountering usability issues during the main sequence.

Issue classification

Level 1 Classification		
Round	Description	Issue
1	Error 1	Grip button is comparatively more difficult to handle (with the trigger button) as more focus is currently laid on the trigger button.
1	Less is more	Simplify the board.
1		Change the game to support a single language.
Level 2 Classification		
0	Failure 3	MRC clip did not help the player notice (the cue on) the AR glasses.
Level 4 Classification		
0	Error 4	Player was not able to distinguish between AR and VR technologies during the interview.

As the solution to the level 3 issue, in terms of finding it within the tutorial phase, is still unclear, more data needs to be gathered. Solutions to levels 1 and 2 issues are going to be implemented in the revised prototype for the next rounds of RITE.

Improvements

1. Incorporate the grip button during the tutorial in the same manner as with the trigger button; ask for the player's input after explanation of the button.

The model of the board, containing the three frames, can be reverted back to the original board model to support a single language during gameplay. Furthermore, the script *ShowMessageOnWallOfTextAction.cs* (figure 42) needs to be adjusted as it currently only supports dual language gameplay.

2. Further simplification: revert board back to original model in order to support a single language during gameplay

The MR clip conveying the AR layer has been updated, incorporating also a first-person view cutscene. The cutscene is added to emphasize the glasses, which after wearing, will activate the AR layer. The perspective has furthermore improved to include not only the AR glasses, but also the entire player. The new clips are visualized in black and the old in light blue in figure 53 on the next page.

3. Implement updated MR clip into the tutorial in order to convey the importance of AR (component of storyline) during gameplay.

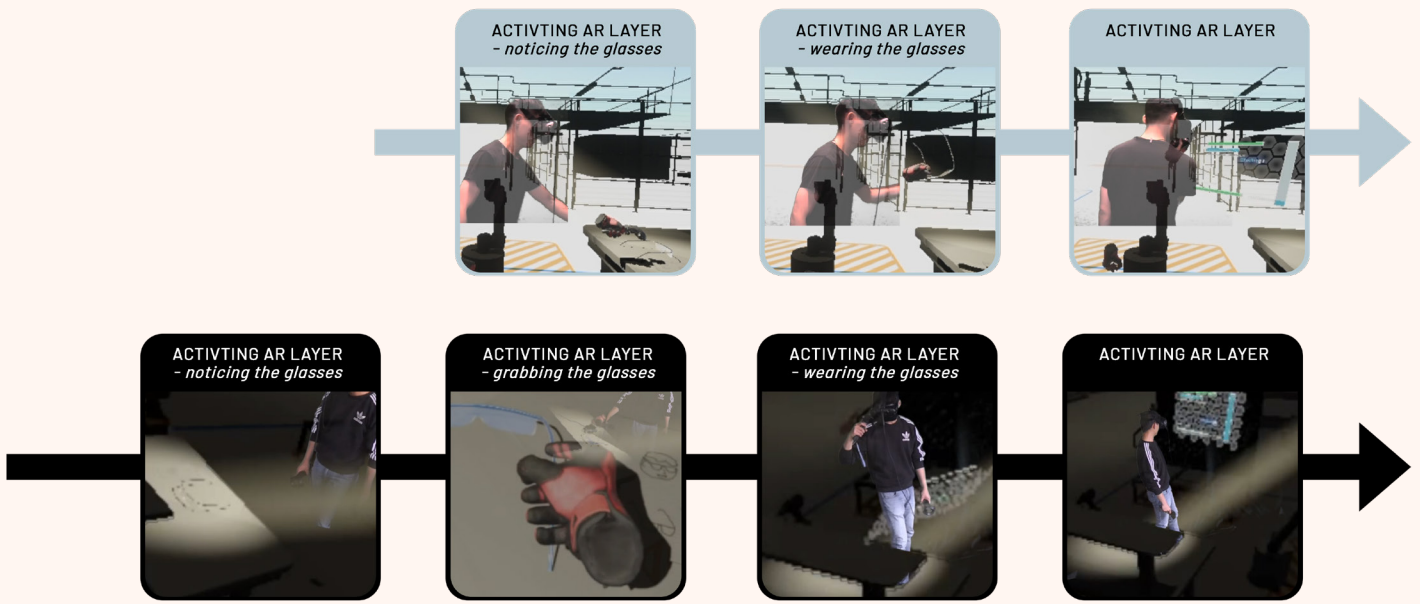


FIGURE 53. UPDATED MIXED REALITY CLIP ABOUT AR (BLACK)

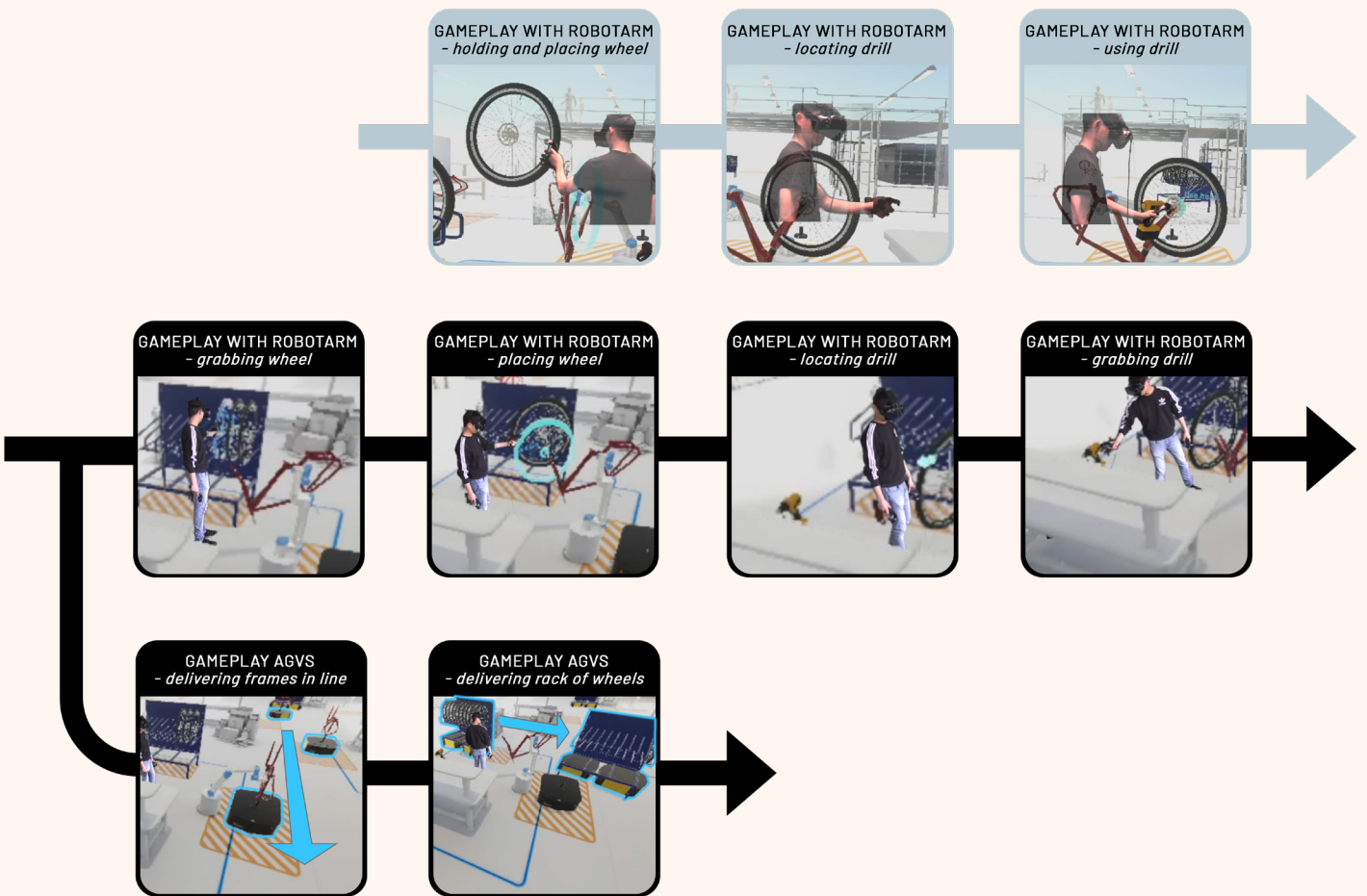


FIGURE 54. UPDATED MIXED REALITY CLIP ABOUT COLLABORATING WITH VARIOUS ROBOTS (BLACK)



The earlier MR clips conveying the gameplay in the main sequence (light blue in figure 54) were relatively more focussed on collaborating with the robot arm; picking and placing the wheel, and also attaching the wheel using the drill. Because the newly added sprites, together with the added messages, explain the basic interaction controls sufficiently and participants fail to (actively) notice the collaboration with robots in the gameplay; the focus in the new MR clips (black in figure 54) has been shifted to collaborative robots in general; including also the supportive work of the AGVs. New clips include, like the earlier reshoot, a zoomed-out perspective; including more of the scene and the entire player into the frame. In addition to collaborating with the robot arm (which will be played back at a higher speed), shots from the AGVs bringing the frames and replacing the rack with wheels are also added (last row in figure 54).

- 4. Implement updated MR clip into the tutorial in order to convey the collaboration with robot component of the storyline to the player and set the context.

The new MR clips will be displayed on the main board, along with short messages to provide narration to the clips. The texts shown during this phase will appear in the green rectangle, situated above the filled out green rectangle, where the clips will be shown. The areas for texts and media (clips or sprites) change when sprites are used (to explain the basic interaction controls). During this phase, sprites will be displayed in the filled out blue square, supported by a relatively larger text field (blue) for clearer explanation. The improved board can be seen in figure 55

- 5. Add two sizes of text fields to the board model (smaller green and bigger blue outline) in order to support the different sizes of media used (video in green and sprites in blue).

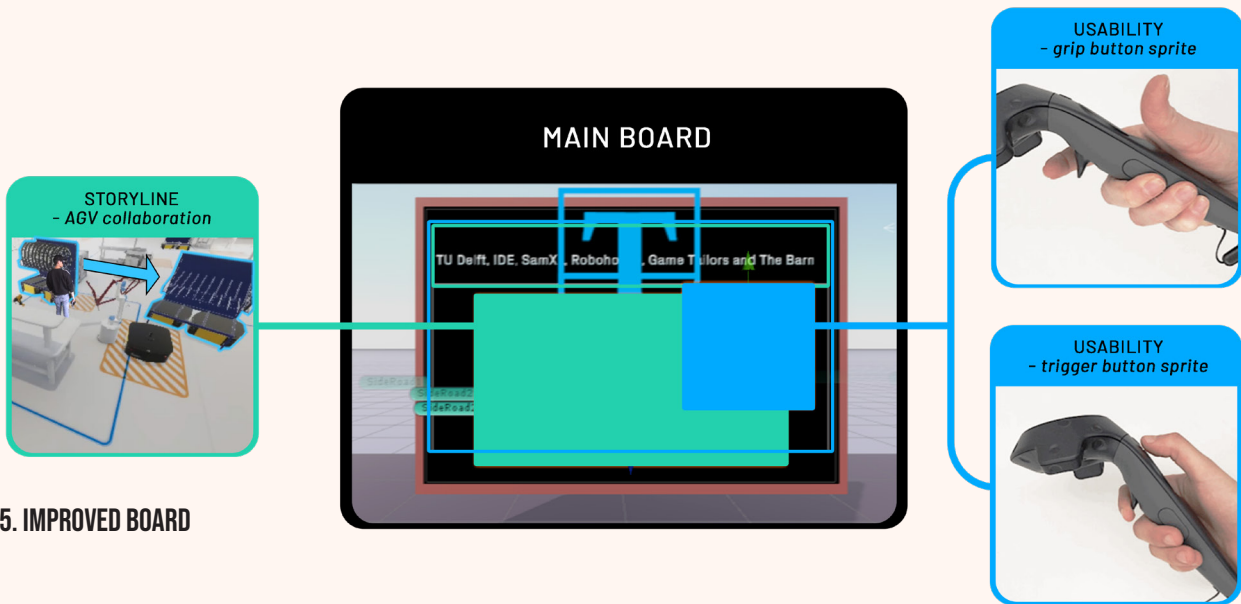
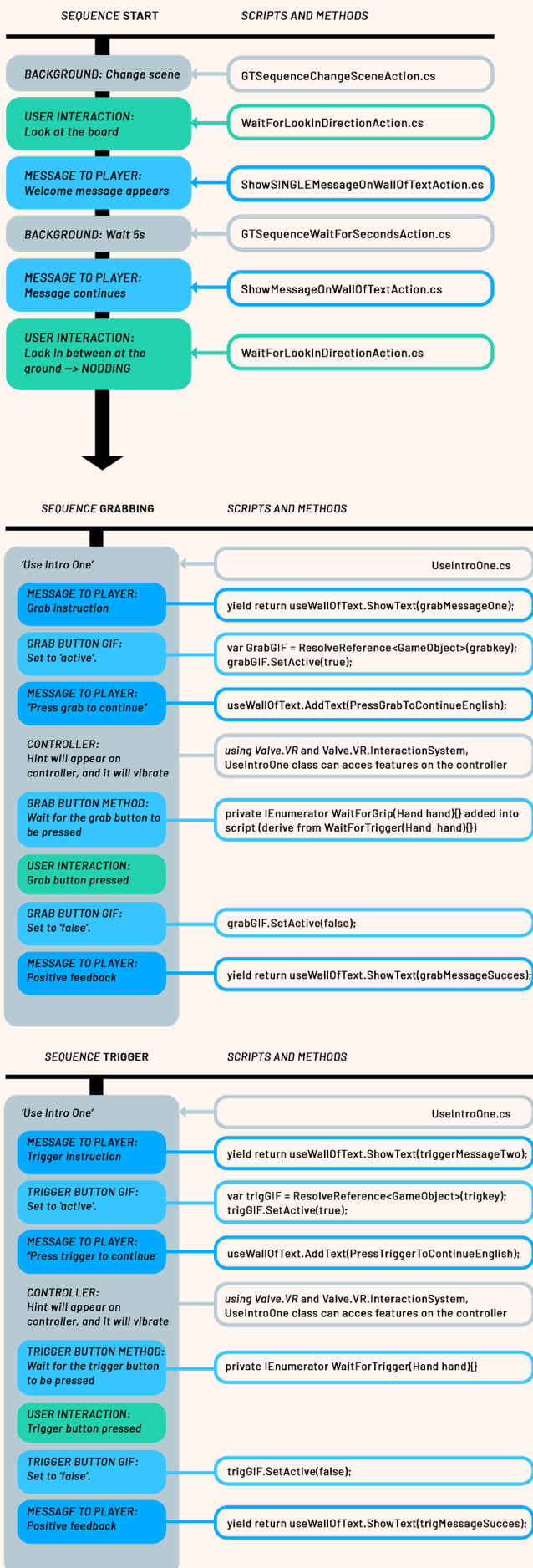


FIGURE 55. IMPROVED BOARD



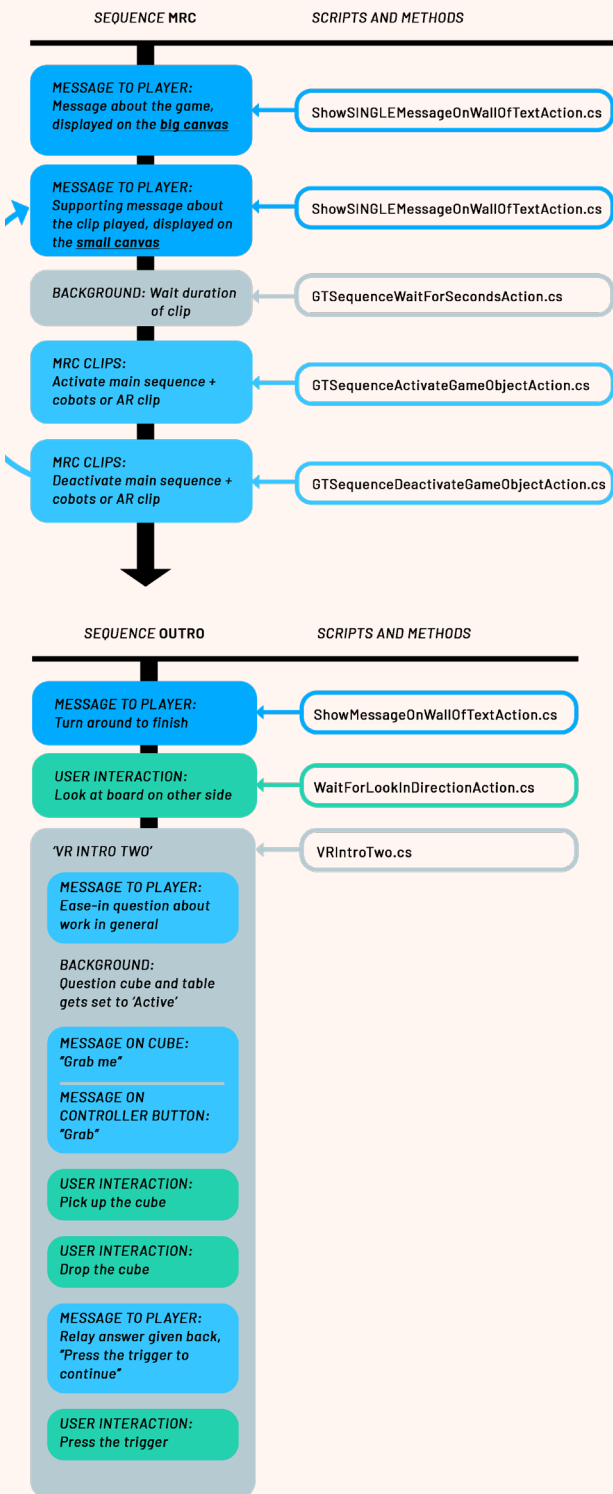
Sequencing intro

The sequencing has been changed much after implementing the improvements. Like the prior version, the basic interaction controls sequence will start after the player nods. However, to make sure the player is engaged with the message conveyed, the intro message will only start once the player looks at the board. This sequence is built using assets derived from individual scripts.

Basic interaction tutorial

In order to access Valve methods for the controller, a new script has been made. The script incorporates two sequences to explain the grab and trigger button. For the grab button, a new method to wait for the grab button input has been added. This method is derived from the existing WaitForTrigger.cs method. Like the prior version, the grabbing (or the trigger) sprite will be supported by a message, appearing within the blue outline in figure 55. When the WaitForGrip(Hand hand){} method has been satisfied, positive feedback will be given to the player.

The sequence for the trigger button is, like earlier mentioned, very similar. The only differences are the media used, supporting texts and the WaitForTrigger(Hand hand){} method. Like the sequence for grabbing, the player will receive positive feedback once the trigger has been pressed.



Mixed reality clips

Like the intro, the sequence conveying the storyline using MR clips is build using individual assets derived from scripts. After the trigger button has been pressed, a bridging message will be displayed, followed by a se-quence of displaying a message (on the smaller green text area in figure 55) alongside a video of varying lengths. This sequence will be repeated three times for the cobot and gameplay clips and the AR clip.

Outro

When the final clip is finished, a message will be displayed (largely, in the blue text area of figure 55) to the player and an identical board will appear behind the player. The question part of the original outro sequence, written in IntroAction.cs, has been rewritten into VRIntroTwo.cs to exclude the prior interactive part of the original script. The sequence will follow the same structure once the player has turns 180° and faces the (newly appeared) board.

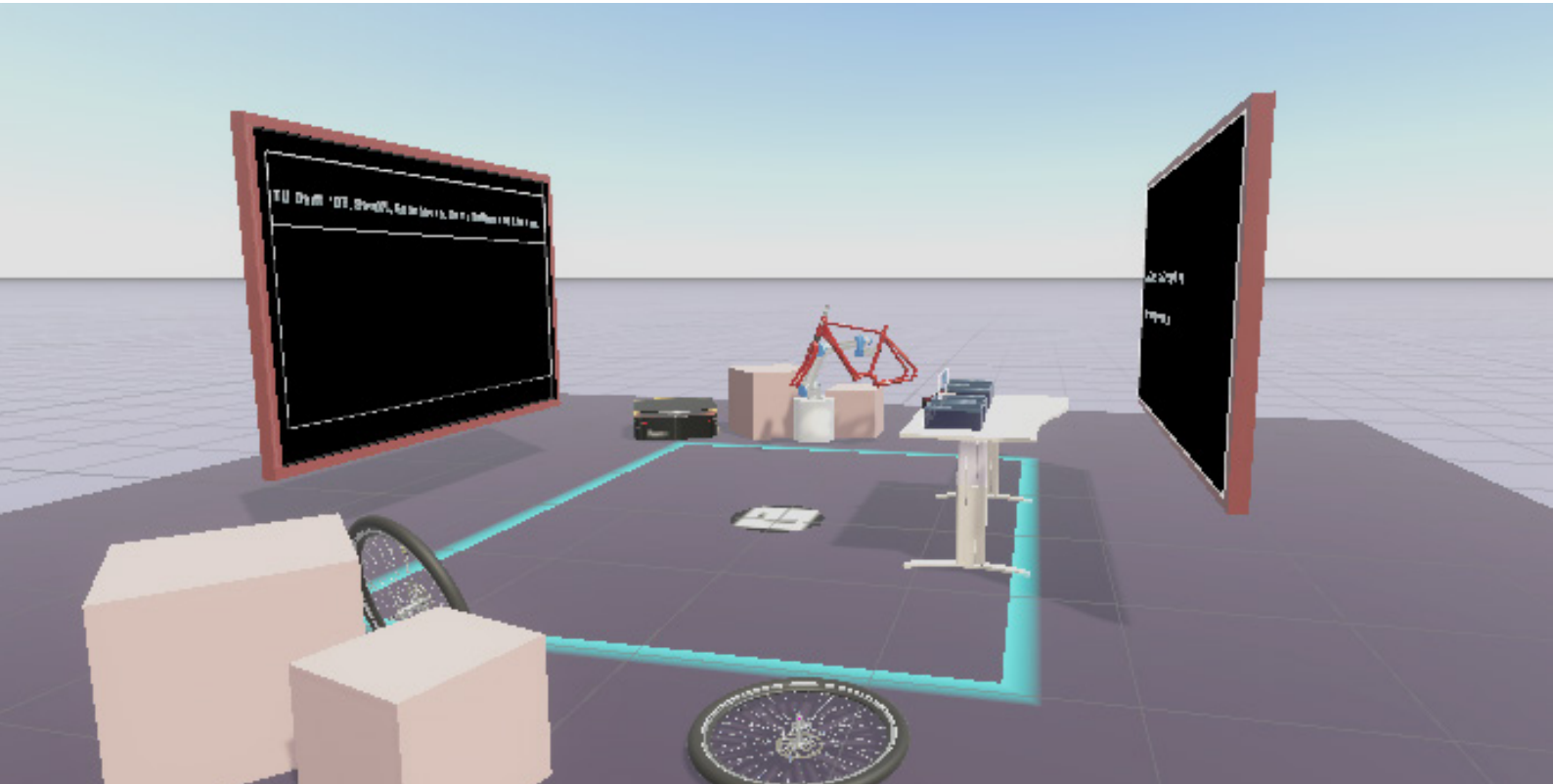


FIGURE 56. FURTHER SIMPLIFIED PROTOTYPE V1.2

PROTOTYPE 1.2: RITE ROUND 2 & 3

The scene of this round looks similar to how the tutorial phase initially looked. However, the board in front of the player contains different sized text areas and includes several media (MR clips and sprites). Although the sequencing has been changed much, the trigger to start and end the tutorial, nodding and turning around (to the elements appearing at the back of the player) respectively, has been kept like the previous round(s). The further simplified prototype can be seen in figure 56.

No notable failures occurred during rounds 2 and 3 however, errors still took place. During the second round, notably, the participant was able to recognize the 'VR layer' activated after wearing the glasses (she meant the AR layer), but was not able to mention the collaborative work with robots during the interview. The collaboration with the robots in the game seem to be not actively noticed, this was also true for the participant during the third round of RITE. Furthermore, like the participant in the pilot round, the participant in the second round was confused between the terms AR and VR.

1. Collaboration with robots and the future of work seems to be not explicitly comprehended.

During the third round, when the end of the tutorial phase is to be triggered (by turning around), the participant only looked over his shoulder. As the trigger needs the player to face the (newly appeared) board at the back of the player, the ending of the tutorial did not start. The participant did not need external help however to solve the error.

2. The text "turn around to end this tutorial" is not explicit enough as some players are not able to trigger the ending.

Like the participant in round 1 of RITE, the connection between the content of the game and the future of work, was only established after further inquiring the participants. Specifically, asking participants to explain how they think bicycles get assembled currently as compared to what they experienced in the game was helpful. The participant in the second round of RITE gave the experience a score of 90 using the SUS questionnaire, comfortably passing the 80.3 mark like the previous iteration. The participant in the third round of RITE gave the experience a system usability score of 80. Less than the previous round, but still within the range of "people love your system and would recommend it to their friends" (80.3+).



Issue classification

Level 1 Classification		
Round	Description	Issue
3	Error 2	The text “turn around to end this tutorial” is not explicit enough as some players are not able to trigger the ending.
Level 3 Classification		
2	Error 1	Difference of the scene played vs how assembly works now needs to be emphasized after gameplay, preferably in the POSTGAME
2 & 3	Error 1	The storyline of collaborating with robots hasn’t been conveyed (effectively)
Level 4 Classification		
0 & 2	Error 4	Player was not able to distinguish between AR and VR technologies during the interview.

As the solution to the levels 3 and 4 issues, in terms of finding it within the tutorial phase, is still unclear, more data needs to be gathered.

Improvements

The message displayed during the outro of the tutorial (to finish the sequence) needs to be more explicit, as some players are not able to trigger the ending.

1. Change the text and make it more explicit when finishing the tutorial (figure 57).

Furthermore, the sprites shown on the board are also added next to the controller. When the game prompts the user for button-input, the controllers vibrates and signs appear; having the sprites appear next to the controllers also ensures players are able to see the hints, regardless of whether they are looking at the board or at their hands/ controllers in VR.

2. Add sprites to controller als (in addition to displaying them on the board)



FIGURE 57. IMPROVED MESSAGE



FIGURE 58. SPRITES NEXT TO CONTROLLERS

PROTOTYPE 1.3: RITE ROUND 4, 5, 6 & 7

The scene of these RITE rounds looks exactly the same as the prototype 1.2; with the only difference being the adjusted message in order to finish the tutorial and the added sprites appearing next to the controllers when basic interaction controls are explained (figure 58).

No failures occurred during these rounds; however, each participant had some errors that could be optimized. It is doubtful to what degree the future of work content is conveyed consistently. Some students, who are familiar with the concept of Industry 4.0, would recognize the future scenario in the game after playthrough (i.e., participant 6). Others, like participants 4 and 5, are not able to mention (specifically) the use of different robots during collaboration. On the contrary, the use of AR in the game seems to be comprehended, as notable failure 3 hasn't occurred since prototype 1.2.

1. The storyline component of collaborating with robots is not conveyed effectively.

Furthermore, some participants (specifically participant 6, who expressed the video to be slow) would try to assert some influence/control over the video's during playback. Participants would nod and press the trigger button; presumably to skip the videos and speed up the tutorial.

2. Players need more elements of control when the MR clips are shown.

Also, participant 7 had some troubles triggering the tutorial through nodding (but did not need external intervention).

3. The trigger to start the tutorial might need some sensitivity adjustments

Generally, most participants mention usability to be intuitive (specifically participants 4, 5 and 6) and the explanation for controls to be sufficient. As usability becomes less of a hurdle, the prototype; which only includes the assembly of one single wheel (as opposed to three in the full game), seems to be too short. For instance, participant 4 ignored the final task to finish the main sequence so he could prolong the experience. Likewise, other participants expressed how they would've liked the game (main sequence) to be longer. Furthermore, as basic interaction controls become intuitive, players tend to get ahead of the game more often (specifically participant 7).

4. Messages could be conveyed faster as participants comprehend the controls and are eager to start.

Issue classification

Level 1 Classification		
Round	Description	Issue
7	Error 3	The sensitivity of the nodding trigger might need to be adjusted so it triggers the next action more intuitively.
Level 2 Classification		
6	Error 2	Players give user input when the MRC are being shown. There seems to be an element of <i>control</i> missing.
Level 3 Classification		
2	Error 1	Difference of the scene played vs how assembly works now needs to be emphasized after gameplay, preferably in the POSTGAME
2, 3, 4 & 5	Error 1	The storyline of collaborating with robots hasn't been conveyed (effectively)
Level 4 Classification		
0 & 2	Error 4	Player was not able to distinguish between AR and VR technologies during the interview.
7	Error 4	Messages could be conveyed faster as participants comprehend the controls and are eager to start.

Participant 7 is the only participant during the last six rounds of RITE (since the implementation of the nodding trigger) to experience troubles with the trigger. More data is needed in order to evaluate the need for adjustment, as the participant did not need external intervention to overcome the error. Likewise, for the level 4 issues, as its solutions is expected to be found outside of the tutorial; more data is needed to propose for and evaluate fixes.

Currently, regarding the collaboration with robots, one MR clip (figure 54) together with one message is used. As currently 5/7 participants are not able to point out the collaboration with several robots during gameplay in the interview, more elements are likely needed in order to convey this component of the storyline more effectively. The solution for this level 3 issue is being worked on and is going to be implemented as soon as it's ready (i.e., in prototype 1.5).

Improvements

A visual timer element is going to be added to the MR clips to convey the duration of the clip. Because players need to view the content in order to set the context and storyline, it is undesirable to allow players to skip the clip.

1. Add a timer feature for players to gauge the length of the MR clips (figure 59).

Additionally, a trigger input is added to the sequence in order to proceed to the next clip and text. This allows all players to read the story at their own pace, while allowing quick readers to proceed after the correlating clips has been shown.

2. Use the trigger button input, in conjunction with the wait for seconds method, to allow players to proceed in the sequence at their own pace.

The delivery of the complete storyline (collaboration with several robots with the help of AR), seems to benefit from a comparison with how bicycles are currently assembled. A reflection of the gameplay, displayed through a MR clip, together with a comparison will be added in the *postgame*.

3. Add a sequence, similar to the sequence for explaining the storyline during the pregame, to the postgame in order to emphasize the future of work content of the game (figure 60).

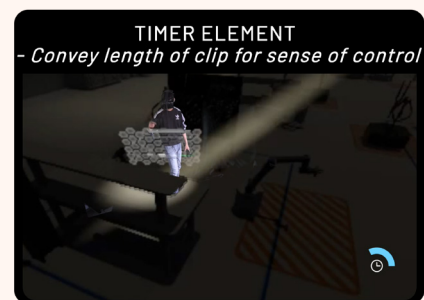


FIGURE 59. TIMER FEATURE

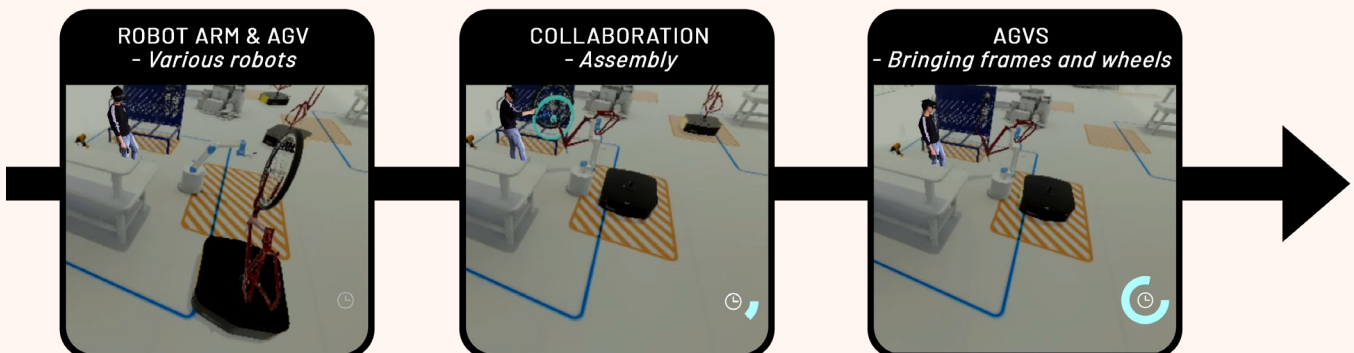


FIGURE 60. POSTGAME (FUTURE OF) WORK COMPARISON

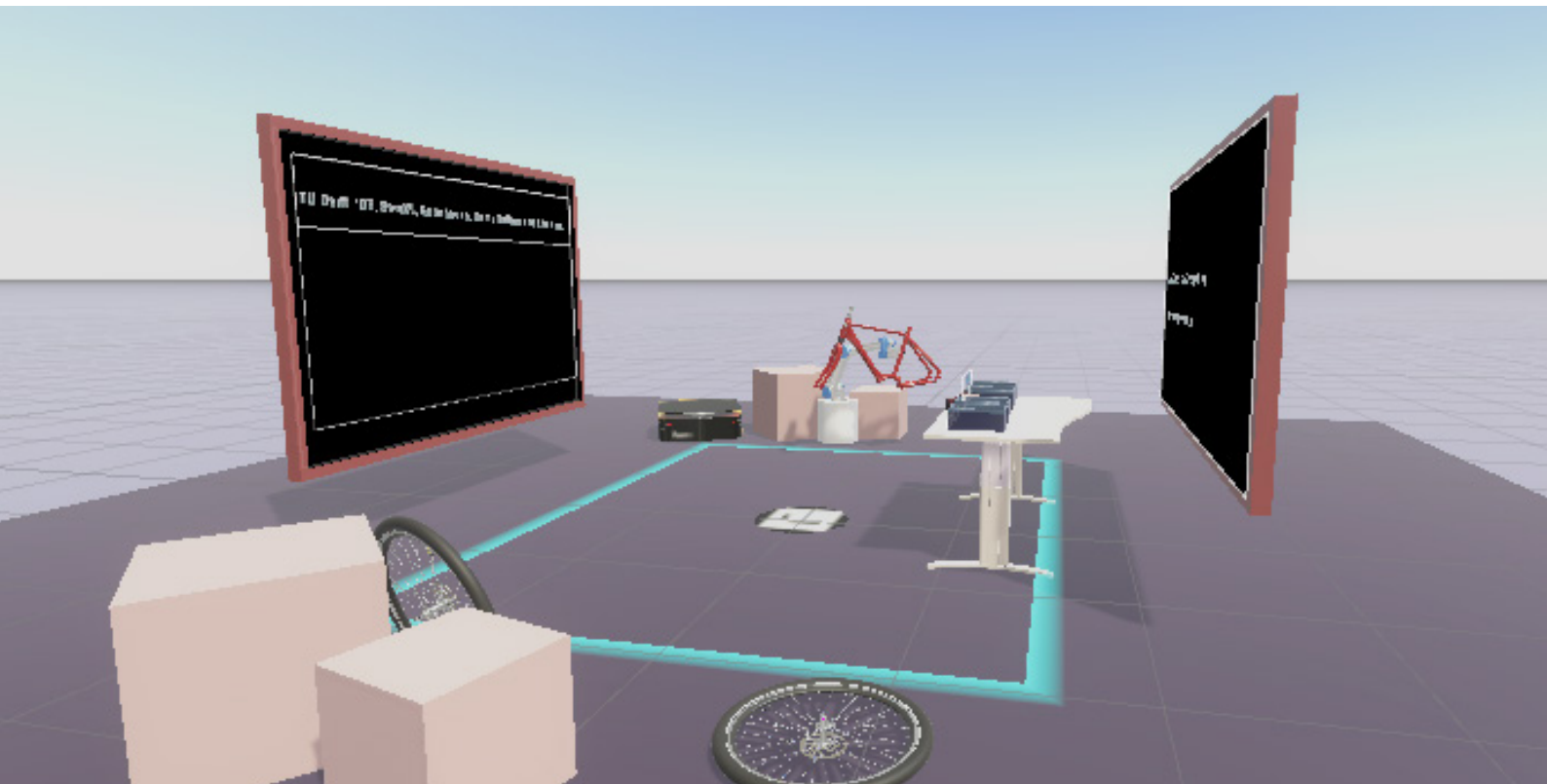


FIGURE 61. PROTOTYPE V1.4

PROTOTYPE 1.4: RITE ROUND 8 & 9

As not much has changed in the scene since version 1.2, the scene in version 1.4 looks identical. However, since version 1.2, much has changed in the sequencing and the media used. Specifically, after the basic interaction controls are explained, MR clips shown have an added timer to convey the duration of the clips. Additionally, as the controls are already explained, texts shown can be skipped using the trigger after the MR clips end. This allows readers of all levels to go through the explanation of the storyline at their own pace. Furthermore, in order to convey the storyline more effectively, a comparison with the current manner of assembling bicycles is made using texts and another MR clip in the postgame.

Like the previous RITE rounds, where version 1.3 was used, only errors occurred. Participant 8 was, like participant 7, eager to start after finishing the tutorial. The messaging during the main sequence of the game was a bit too slow, as after placing the wheel the participant was already moving onto to another wheel (while the drill needed to be used). Also, participant 9 missed a little bit of the MR clip in the postgame as it started playing without the user paying attention.

1. The player's attention needs to be grabbed before starting the sequence added in the postgame.

Nonetheless, both participants were able to mention the collaboration with robots and the help of AR during the interview.

Issue classification

Level 1 Classification		
Round	Description	Issue
9	Error 1	Participant missed a small part of the outro as it started without the player paying attention
Level 3 Classification		
2, 3, 4 & 5	Error 1	The storyline of collaborating with robots hasn't been conveyed (effectively)
Level 4 Classification		
0 & 2	Error 4	Player was not able to distinguish between AR and VR technologies during the interview.
7 & 8	Error 4	Messages could be conveyed faster as participants comprehend the controls and are eager to start.

Improvements

The sequence of the outro in the postgame starts when the player looks at the board. A trigger will be added to establish that.

1. Add trigger to start the sequence (including the MR clip and messages) during the postgame

Furthermore, the explanation of collaborating with robots has been expanded using additional messages and the AGV and robot arm models. The sequence to convey the storyline has additionally been relocated to the board at the back of the player; replacing the original question posed together with the table (with cube and four containers).

2. Add the use of models and message to support the collaboration with robots component of story. Relocate the sequence to the back of the player, after the basic user controls are explained.

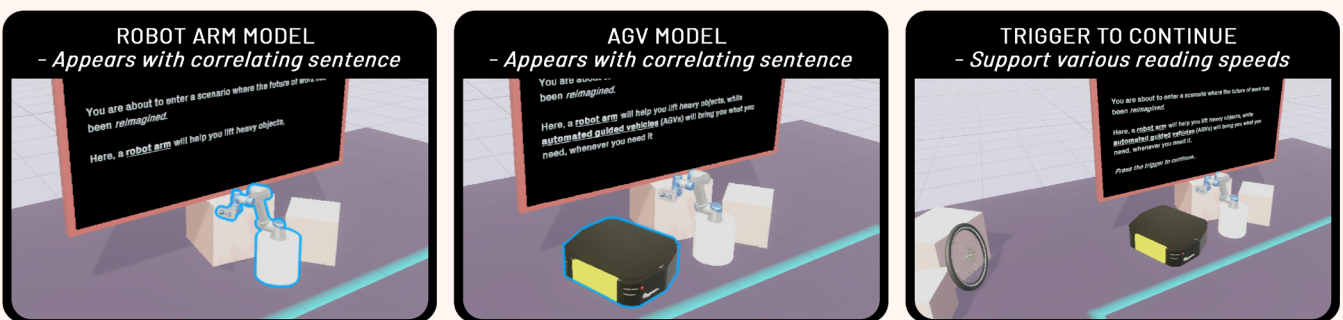


FIGURE 62. EXTRA SEQUENCE EXPLANATION FUTURE OF WORK COMPONENT USING 3D MODELS

As no failures has occurred since prototype version 1.2 and little grab and trigger interactivity currently exists (the table with the cube and container has been replaced/removed), the sequence to explain the basic interaction controls can benefit from more interactivity. Using the same table (model), together with the drill used in prototype 1.0 (during the pilot round), more grab and trigger interactivity is added to the sequence. The order of the elements used can be seen in figure 63.

3. Add more interactivity while explaining basic interaction controls.

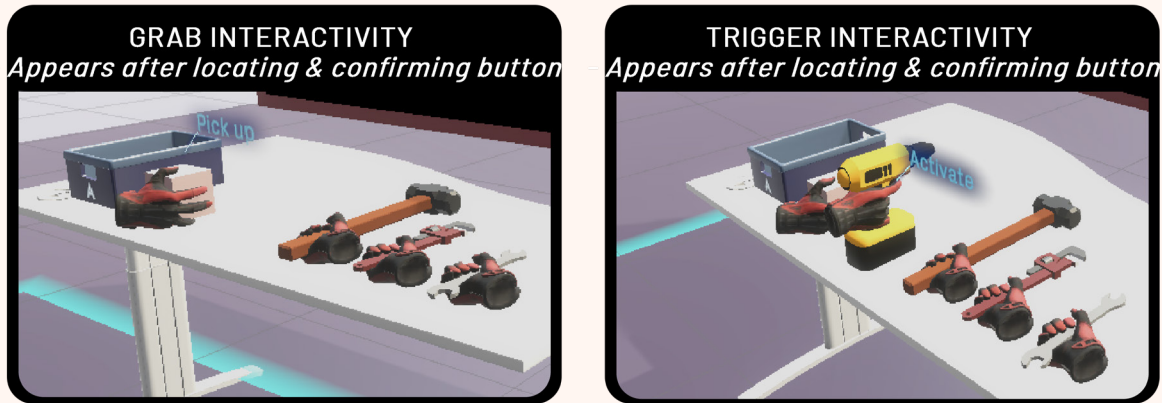


FIGURE 63. UPDATED MIXED REALITY CLIP ABOUT AR (BLACK)

Furthermore, as little attention is given to the physical controllers, a brief moment is added for players to find and identify the grab and trigger buttons before starting the basic interaction controls sequence.

4. Add a moment where players are able to identify & locate the needed buttons on the (physical) controllers.



FIGURE 64. BUTTON LOCALIZATION

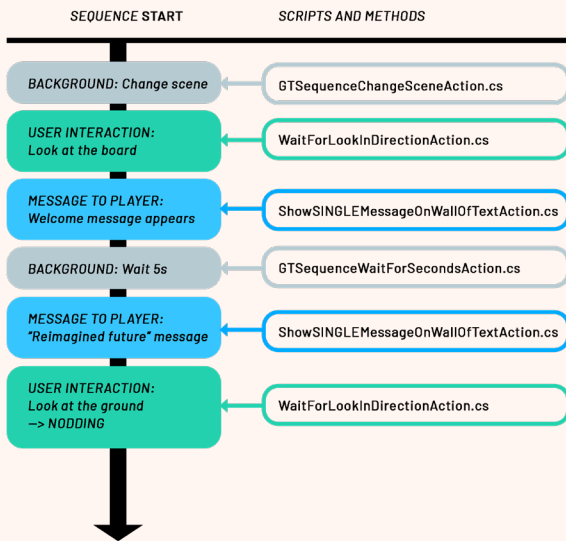


FIGURE 65. INTRODUCTION/START SEQUENCE FLOW

Sequencing

intro

Players are welcomed to the experience when they look at the board. The context of "futures" is set after with an additional message and the localization sequence (figure 64) start when the player nods. the flow can be seen in figure 65.

Localization sequence

Because players usually are already wearing the HMD (physically) when the controllers are passed, they do not have the chance to take a good look at the controllers and its buttons. After the player nods, the player is prompted in VR to look at the controllers, where signs will appear, highlighting the relevant buttons for the experience.

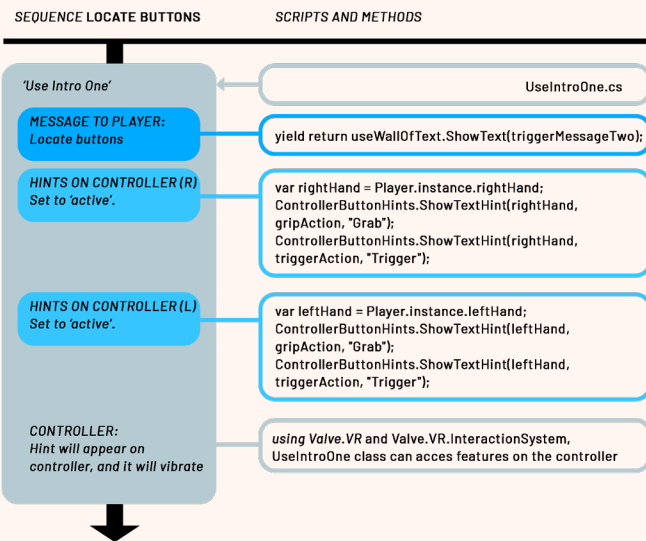


FIGURE 66. LOCALIZATION SEQUENCE FLOW

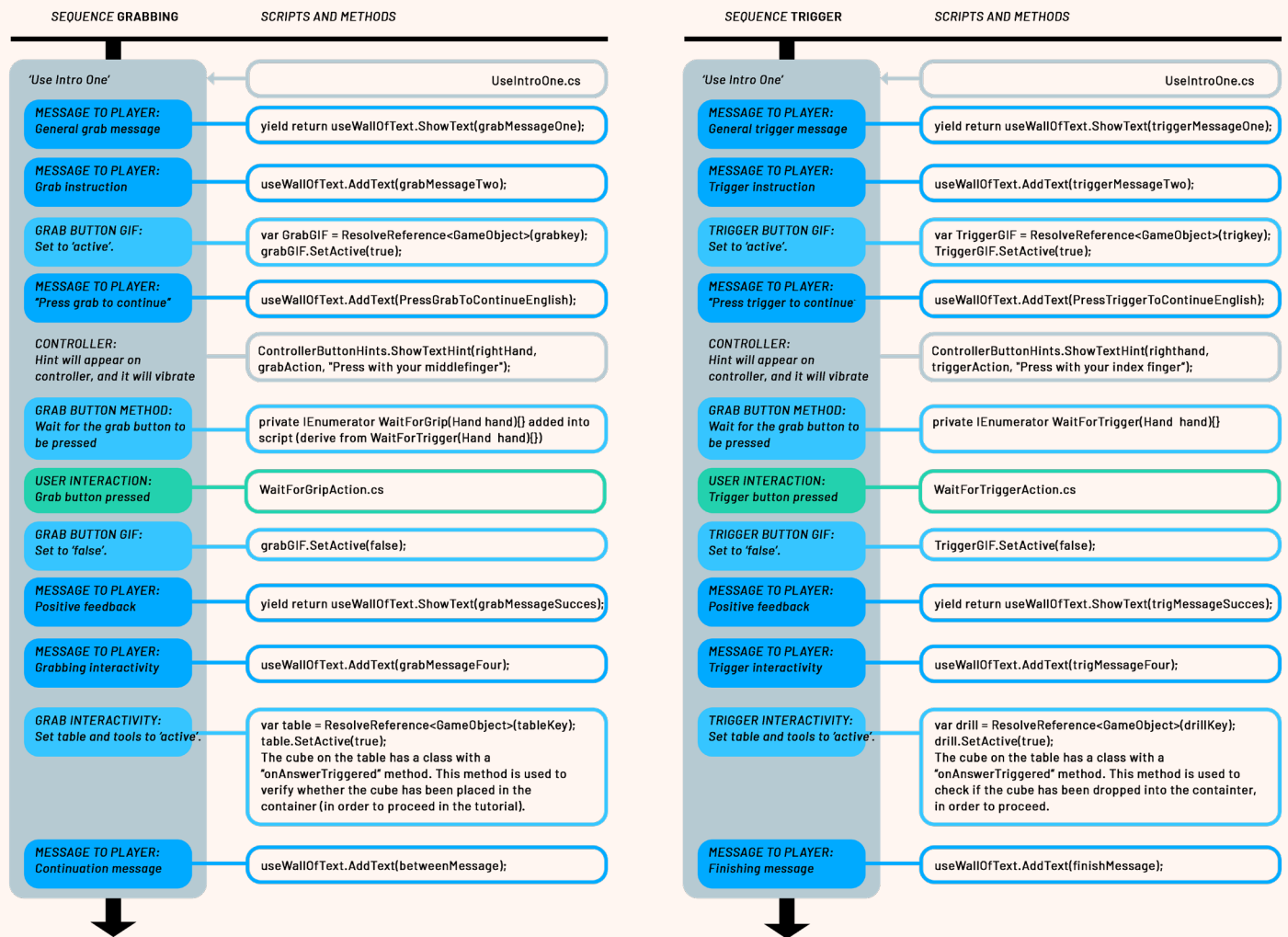


FIGURE 67. GRABBING AND TRIGGER SEQUENCE FLOW

Grabbing sequence

The first three messages form one explanation that gets slowly build up. When the player is prompted to push the grab button, sprites appear on the board and next to the controllers (figure 58). Positive feedback is given when the button is pressed and the desk containing the cube and tools appears in front of the player after (figure 63). Players are able to grab and play with the tools if desired and can end this part by placing the cube in the container.

Trigger sequence

The trigger sequence starts when the cube is placed in the container and is very similar to the grabbing sequence. The only difference being the center of focus (trigger button as opposed to grab button) and media used. When the trigger button is pressed, a drill will appear in between the cube and the earlier appeared tools (figure 63). A message prompts players to pick up and activate the drill. Like the grabbing sequence, the cube needs to be put into the container to finish this part of the tutorial. When the cube is placed, the desk gets deactivated and a board appears behind the player. A finishing message is shown to the player, asking him/her to turn around and face the other direction. The sequence flow is depicted in figure 67.

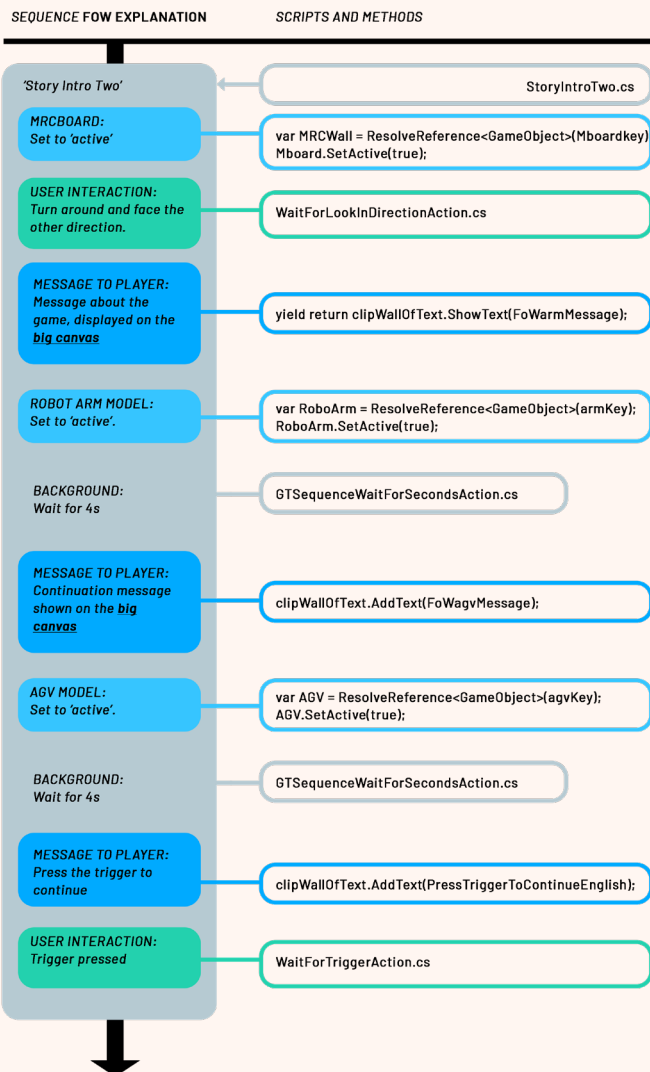


FIGURE 68. FOW SEQUENCE FLOW

FoW and MR clips sequence

The collaboration with various robots is introduced using three messages and two 3D models, building one explanation (figure 62). After the the player presses the trigger, the sequence of MR clips starts as depicted in figures 53 and 54.

During this sequence, players need to watch all clips, but can continue by pressing the trigger after they are shown. This allows readers of all levels to comfortably get through.

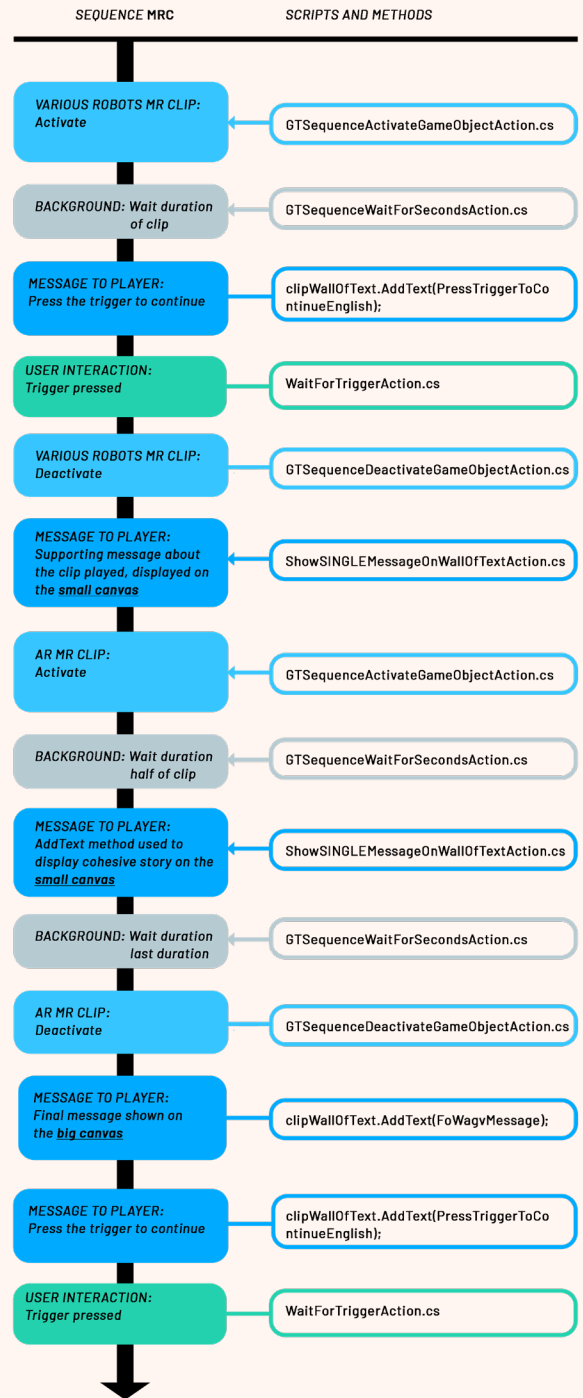


FIGURE 69. MR CLIPS SEQUENCE FLOW

PROTOTYPE 1.5: RITE ROUND 10 & 11

The scene looks similar to the scene used in earlier prototypes. However, there are distinct differences as the sequencing has been changed much to incorporate more interactivity. In earlier iterations, a table with a cube and four containers was located in front of the board, at the back of the player. In version 1.5, 3D models of the robot arm and AGVs can be found in its place instead, supporting the explanation of the storyline. Additionally, the table has been relocated to the board in front of the player. Furthermore, instead of four containers and a cube, the table now contains one cube and one container, together with a drill and several other tools to support the explanation of basic interaction controls and add more interactivity.

Unsurprisingly, similar to earlier rounds, no failures occurred. Furthermore, during the interviews, both participants were able to mention the collaboration with robots and the use of AR. However, like some participants in earlier RITE rounds, participant 11 was not able to differentiate between AR and VR technologies. Also, during the playthrough of participant 10, the left controller temporarily lost connection. Although no external intervention was needed and no notable errors were made, it did cause some confusion for the player as the left hand suddenly disappeared (from the perspective of the player).

Issue classification

Level 4 Classification		
Round	Description	Issue
0, 2 & 11	Error 4	Player was not able to distinguish between AR and VR technologies during the interview.
7 & 8	Error 4	Messages could be conveyed faster as participants comprehend the controls and are eager to start.

Improvements

In earlier iterations, the table (together with the interactions with the cube) was used to communicate the boundaries of the play area. As the objects would appear behind the player, at the edge of the play area, participants were encouraged to physically turn around and walk. Because the objects have been changed and relocated to serve another purpose (adding interactivity during the explanation of basic interaction controls), a gap has appeared in conveying the (physical) limits of the play area.

A small sequence is added before explaining the basic interaction controls in order to bridge this gap. Here, messages will emphasize the blue outlined area on the ground and the 'mesh wall' (Valve assets used in the script call this feature the 'chaperone') players see in VR, in order to communicate the bounds of the play area. The chaperone will temporarily be forced visible (method of the chaperone class) for this.



PROTOTYPE 1.6: RITE ROUND 12, 13, 14 & 15

Version 1.6 includes explanations of the VR bounds and basic interaction controls. It also sets the context of the game-content (a specific 'potential plausible' future scenario; revolving around collaborating with robots with the help of AR) by making use of several MR clips players get to see before and after playing the game (i.e., in the pre- and postgame).

Four rounds of evaluations were done using this prototype, in which only one error occurred and no failures. Furthermore, all participants were able to mention components of the future of work during the interview. Before starting the main sequence, in the factory, participant 12 was confused for a brief moment as the light switch (that needed to be turned on) was located slightly outside of the mesh wall. Also, the participant noted not to know which controller to use during the tutorial.

1. Specify which controller to use in the tutorial (texts) to avoid possible confusion.

Additionally, participants 12 and 15 advised the MR clips to be displayed slightly bigger, as they currently did not feel fully involved. Nevertheless, during the interview, both participants were able to describe the future of work storyline.

Issue classification

Level 1 Classification		
Round	Description	Issue
12	Error 1	Specify which controller to use in the tutorial (texts) to avoid possible confusion.
Level 4 Classification		
0, 2 & 11	Error 4	Player was not able to distinguish between AR and VR technologies during the interview.
7 & 8	Error 4	Messages could be conveyed faster as participants comprehend the controls and are eager to start.

Improvements

During the explanation of the basic interaction controls, the right controller will be specified in the messages when asking for user input.

RECAP

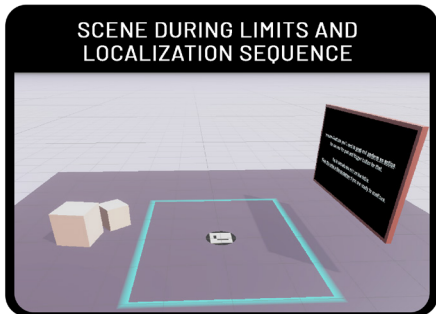


FIGURE 70. SCENE: INTRO SEQUENCE

In version 1.6 of the tutorial, players get eased into the context of (plausible) futures following the sequence as described in figures 65 and 66. However, in between both flows, before the localization of buttons starts (figure 66), a new sequence is added where the bounds of the play area are emphasized. As the chaperone is temporarily (for the duration of this sequence) set to active, the limits of the space is, together with the blue outline on the ground, conveyed to the player. After, following the flow depicted in figure 66, players are able familiarize themselves with the controllers; locating and identifying the relevant grab and trigger buttons (figure 64). Figure 70 shows how the scene looks like when this takes place.

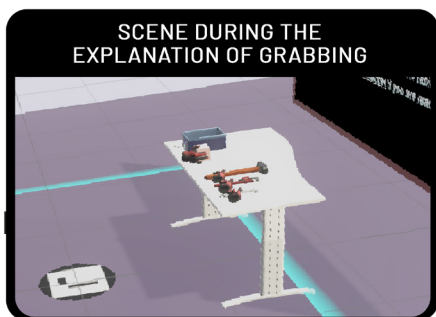


FIGURE 71. SCENE: BASIC INTERACTION CONTROL SEQUENCE

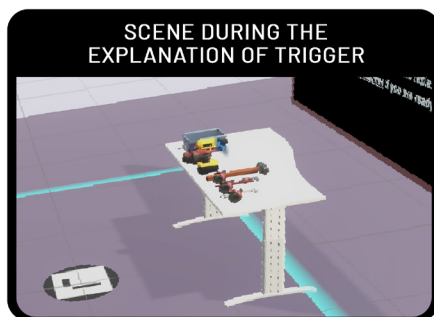


Figure 55 depicts how different media and connected texts are shown on the board. Figure 67 depicts the flow of the basic interaction control explanation. After the button in concern has been pressed initially, guided by the messages and sprites (figure 58), players are able to confirm the action of the button more interactively; as depicted in figures 63 and 71.

The potential plausible future of work scenario, including the collaboration with various robots and the help of AR components, are all separately conveyed. The general future of work storyline is conveyed through a MR clip (figure 60), shown in the postgame (same scene as depicted in figure 70) after the player has experienced the main sequence; in order to stimulate players to think about how currently the same work is done (versus what they experienced in the game). The collaboration with various robots component of the scenario is conveyed using the 3D Models (figure 62) together with a MR clip (figure 54). The support of AR component is conveyed using a MR clip (figure 53). Figure 72 shows how the scene looks like when this takes place. The general flow of the sequence is depicted in figures 68 and 69.

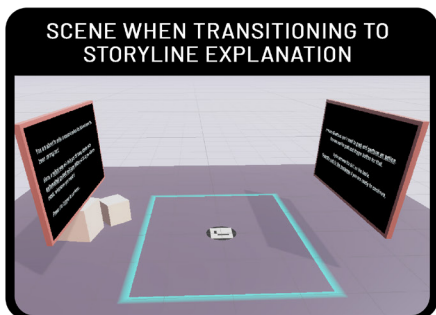


FIGURE 72. SCENE: STORYLINE CONVEYANCE SEQUENCE

Two main jobs to be done were distilled from the desired scenario described in figure 36: convey the Future of Work story, in order to provide for more clarity, enjoyment and have an overall more cognizant experience of the content of the game, and explain the basic interaction controls in order to allow for presence and immersion to take place. For this, the tutorial has been redeveloped with the help of a brainstorm, including also input and fixes proposed by participants from the preliminary study.

As the jobs to be done are very subjective in nature, both being very dependent on user-specific characteristics; the specifics of the fixes proposed have been tailored towards (a subset of) the envisioned target group (figures 37 and 46) through the RITE method. The goal of using the RITE method was to implement and validate fixes in order to achieve the subjective jobs to be done as quickly as possible. As the method includes the involvement of relevant user during the (re)development, and fixes are refined and validated according to user evaluations, the (re) development phase satisfies all criteria (as described by ISO 9241-210, 2019) for a human-centered approach.

DATA ANALYSIS

Observation

In order to gauge to what degree basic interaction controls are explained effectively, a list of actions all participants need to be able to go through (without exception) has been distilled from the analysis of the preliminary study. Furthermore, the meaning of failures and error were defined sharply in order to classify the issues accordingly. Failures, errors and revisions of the prototype are visualized in figure 73. Descriptions of errors and failures for each round can be seen in table 2. The data from the pilot round is included as the test setup was changed after the pilot to include less data (in order to implement and evaluate fixes as quickly as possible) than initially gathered.

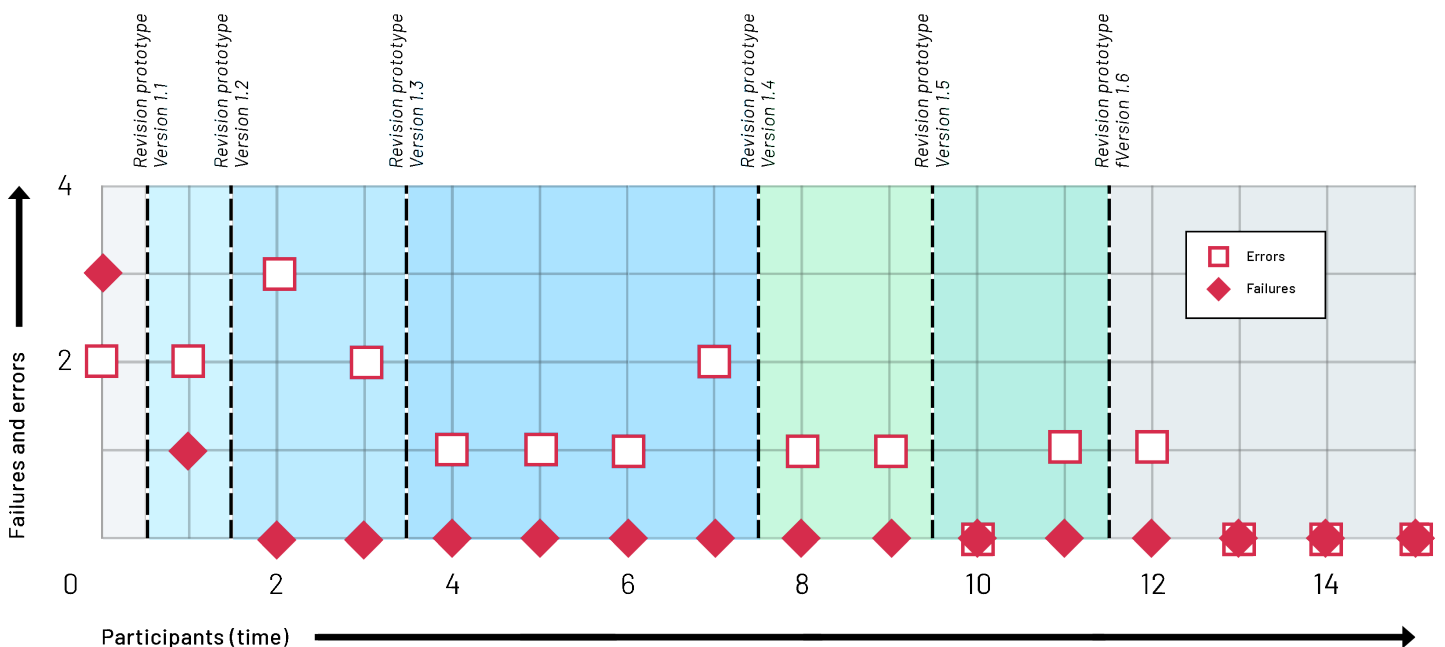


FIGURE 73. FAILURES AND ERRORS DURING RITE

0	Failures	Trigger button failure.	Grip button failure.	AR glasses interaction failure.
	Errors	Sequence unintentionally triggered.	Not able to distinguish between AR and VR technologies.	
1	Failures	AR glasses interaction failure.		
	Errors	Struggled with the grip button.	Not able to distinguish between AR and VR technologies.	
2	Errors	Future of Work storyline not comprehended.	Collaborating with robots storyline component not comprehended.	Not able to distinguish between AR and VR technologies.
3	Errors	Turn around text not specific enough.	Collaborating with robots storyline component not comprehended.	
4	Errors	Collaborating with robots storyline component not comprehended.		
5	Errors	Collaborating with robots storyline component not comprehended.		
6	Errors	Player tries to assert control when the MR clips are being shown.		
7	Errors	Player had troubles with the nodding interaction.	Player was ahead of the game at moments.	
8	Errors	Player was ahead of the game at moments.		
9	Errors	Part of the outro was missed as the clips started playing without the player paying attention.		
10	Errors	-		
11	Errors	Not able to distinguish between AR and VR technologies.		
12	Errors	Player was briefly confused which controller to use during the tutorial (left or right).		
13	Errors	-		
14	Errors	-		
15	Errors	-		

TABLE 2. DESCRIPTIONS OF FAILURES AND ERROR RITE

Failures and errors stated in the list of notable actions have mostly been solved since the introduction of prototype version 1.2 (from the rounds of participant 2 onwards). However, the specifics of fixes and placement of texts and media used went through several iterations in order to make sure the fix had the desired effect (and caused no further issues).

In the desired scenario, described in the synthesis chapter, it is stated that a minimal number of usability errors would occur during the tutorial phase ideally, and zero errors should occur once the player enters the factory. Version 1.5 (from the rounds of participant 10 onwards) included a relatively big revamp of implemented features, including the implementation of 3D models (i.e., robot arm and AGV) when explaining the collaboration with robot component of the storyline and more interactivity when explaining basic interaction controls. Since its introduction, six more participants went through the experience. None of them encountered significant usability errors during the main sequence of the game. As no notable failures and errors has occurred since version 1.2 and no significant errors have been made (also in the main sequence of the game) since version 1.5; the job to be done, explaining basic interaction controls, has been fulfilled (as the desired scenario describes).

Interview

The future of work storyline included two main components, collaborating with robots (1) with the help of augmented reality (2). The comprehension of the AR component could (among the interview data) be derived from whether notable failure 3 (Using the controller to press the button to grab and hold the glasses to put on), specified in the test setup, would occur. As when participants understand the importance of AR in the story of the experience, the spotlight usecue, shining heavily on the AR glasses/desk (communicating the required interaction of putting on the AR glasses), would be comprehended. During the preliminary study, 83% of the participants needed external intervention in order to understand the required interaction. During redevelopment, no external intervention was needed after implementing version 1.2 of the prototype. However, as priorly stated, the last version of the fix was only finalized after implementing version 1.5.

As described in the test procedure, an interview was held after every participant playthrough in order to gauge to what degree the storyline has been conveyed effectively. In contrast to the AR component of the storyline, the storyline component of collaborating with robots doesn't have an (action) indicator (i.e., notable failure 3). It is therefore relatively more difficult to gauge to what degree this component of the story is conveyed organically (without explicitly asking for it). Notably, more features were needed (sequence included the 3D models of the robot arm and AGV) in order to convey the collaboration with robots component of the storyline effectively, compared to the AR component. Additionally, a moment of gameplay reflection, supported by another MR clip; where players would be prompted to compare the work in VR with how the same work is performed currently, was needed in order to effectively convey the future of work storyline in general. After the revamp of features in version 1.5, all participants were able to mention the future of work storyline, including the collaboration with robots using AR, during the interview. However, participants generally needed more inquiring in order to mention all components.

Aside of the area described in the test procedure, the VR corner at the Applied Labs (TU Delft, faculty of IDE) was also used for RITE. Consequently, two types of students have been recruited; university level students (mostly from TU Delft) and intermediate and higher vocational education (MBO and HBO) level students (mostly from the Hague).

During the interviews, 4 out of 15 participants were not aware of the difference between AR and VR technologies, presumably due to differences in education level. Furthermore, some students were relatively more informed about the future of work and were able to mention Industry 4.0 or SAMXL, presumably because they were university students from Delft.

As notable failure 3 (concerning the use of the AR glasses in the experience) does not occur and participants are able to mention the collaboration with robots or the future of work in general during the interviews (since version 1.5), the job to be done of conveying storyline of the game is reasonably fulfilled. However, although the future of work scenario and most of its components are conveyed through the experience, the efficacy is still very dependent on the player. Because some players might not understand the difference between AR and VR, the AR component of the story (conveyed through VR) can be hard to comprehend. Additionally, participants generally needed more than a few questions (i.e., further inquiring) in order to mention all components of the storyline.

System Usability Scale

In order to give an idea of the general usability of the system, based on the participant's self-reporting as opposed to the researcher's observations and notes, participants scored the usability of the experience with the help of the system usability scale. The SUS score was calculated by first subtracting 1 from the uneven questions and subtracting the values from the even questions from 5. The resulting values are summed and multiplied by 2.5 to get the final score. When the resulting score is above 80.3, people generally love your product and would recommend it to others (A). When the score is above 68, the usability is rated as above average; it's okay but could still somewhat improve (C). When the score drops below 51, the usability is rated as below average and needs improvement (F). Scores for every participant are graphed in figure 74.

Notably, the usability was rated slightly below average, with a score of 65, when (re)development started during the pilot round (with 3 failures and 2 errors occurring). During (re)development, aside of one measurement (participant 11), the prototypes generally get scored in between the range of C; usability is okay, but could still somewhat improve and A; people generally love your product and would recommend it to others. The highest scores were given for version 1.6 of the prototype. However, as relatively few people tested one version of a prototype (version 1.6 used 4 participants), no conclusions can be drawn with strong statistical relevance. Nonetheless, on average, all prototypes following version 1.0 scored higher when compared with version 1.0. When it comes to usability, later iterations seem more preferable.

Furthermore, remarkably, even with all the basic interactions control and clarity issues at the start of (re)development, participants only rated the game with an F twice; at the start with version 1.0 and when version 1.5 was implemented (participant 11). Evidently, the use of VR is very desirable, as regardless of whether participants need external intervention in order to get through the experience (the participants using version 1.0 and 1.1 definitely did), the experience generally gets an above average SUS score.

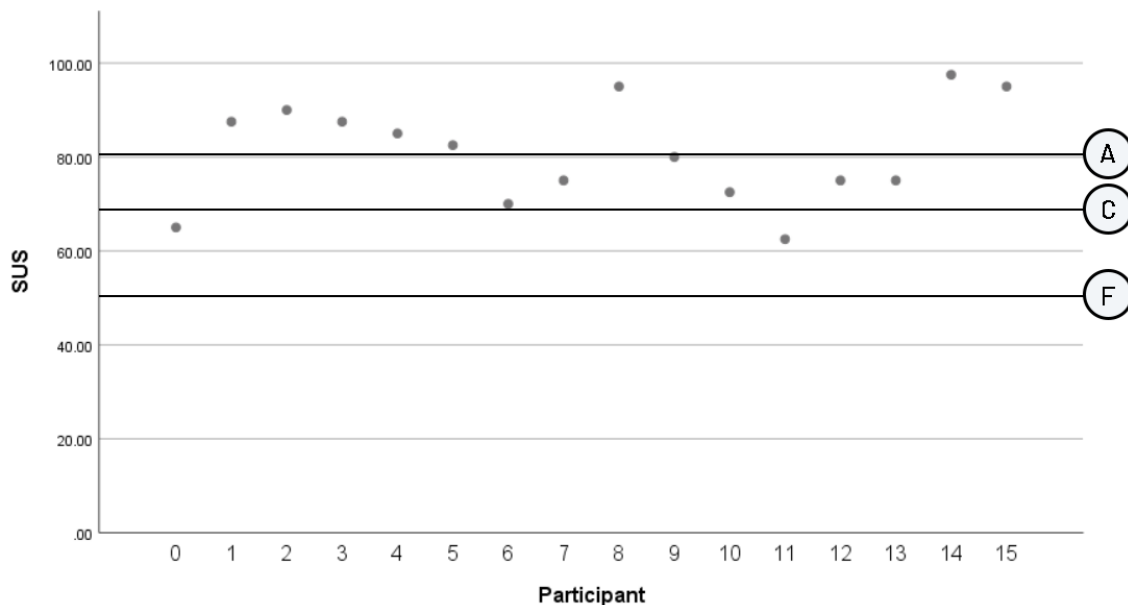


FIGURE 74. SUS SCORE GRAPH



Conclusion

The scenes in version 1.0 and 1.1 were slimmed down considerably as players tend to get distracted by the objects that appear in VR and are placed within (the vicinity of) the play area. Between version 1.2 and 1.5, specifics of the MR clips, sprites, messages and sequencing of the features used, in order to explain the basic interaction controls, underwent several refinements. MR clips have been reshot twice and re-edited three times. Furthermore, texts and triggers were added and refined multiple times in order to achieve its goals. Additionally, a sequence conveying the future of work storyline in general was added to the postgame. As implemented features were evaluated and refined, a revamp of the sequencing, implementing more interactivity during the explanation of basic interaction control, was introduced in version 1.5. The final version, version 1.6, contained small refinements of the revamp implemented in version 1.5

In order to fulfill the jobs to be done of explaining basic interaction controls and explaining the storyline and its components (reasonably effective), the solutions and fixes went through several iterations. The iterations were driven by user evaluations in order to make sure features implemented had its desired effect(s) and didn't cause other issues. As earlier stated, compared to traditional usability testing; the focus on determining whether a fix solves a problem and what the likelihood is of the fix causing other issues through iterations in the RITE method, is how RITE discerns itself from traditional usability testing.

Following the RITE method, also taking into account failures and errors observed in the latest version (1.6) of the prototype, it is concluded that the lack of basic interaction control is fixed. SUS scoring for this version ranged from 75 to 97.5 points; comfortably passing the above average mark of 68 points overall, and occasionally entering the range where people love the product and would recommend it to others (A mark). Although the future of work scenario is generally conveyed to the player using version 1.6 of the prototype, the mention of specific components of the storyline, using AR (1) in order to collaborate with robots (2), needed more inquiring than the initial interview script asked for. Furthermore, to what degree the storyline is conveyed, and how much further inquiring was needed, is very dependent on the pre-existing knowledge of the player. Some, who were familiar with the Industry 4.0 concept, would grasp the storyline and all its components relatively quickly, while other had troubles discerning between AR and VR in general. Because of this, it is concluded that the future of work storyline is reasonably conveyed, but could be improved.

Currently, 20% of the participants were not able to discern between AR and VR. Although enough data has been collected to validate the issue (and possibly lift the classification to level 3), no fix has been proposed nor implemented yet. As AR is a component of the future of work storyline, solving the issue will likely also benefit the effective and complete conveyance of the storyline (and its components). Furthermore, as usability becomes more intuitive, more participants get ahead of the game (approximately 13% during the RITE iterations) due to the messaging in the main sequence of the game being too slow. Because the (re) development was focused on implementing the two jobs to be done into the tutorial, easing participants into the game before experiencing the main sequence, no fix has been implemented.

Level 3 Classification		
Round	Description	Issue
0, 2 & 11	Error 4	Player was not able to distinguish between AR and VR technologies during the interview.
7 & 8	Error 4	Messages could be conveyed faster as participants comprehend the controls and are eager to start.

Conclusion

VIRTUAL REALITY FOR FORESIGHT

Using the experience as foresight effectively, requires the potential plausible future of work scenario to be conveyed effectively. Although the future of work scenario is generally conveyed, the mention of specific components of the scenario (i.e., collaborating with robots and the use of AR) is more difficult. This is mostly caused by differences in pre-existing knowledge of the players, as some of the were aware of the Industry 4.0 concept during the RITE redevelopment, while others were not familiar with (the differences between) AR and VR. Remarkably, SUS scoring has indicated that participants are generally very positively inclined when it comes to the use of VR, regardless of the need of external intervention due to usability issues.

Using VR as tool for foresight is evidently very desirable from the perspective of the user. However, using VR alone, without a conversation after (or another way to spark thoughts), is not always sufficient for the conveyance of the complete scenario (including the collaboration with robots and AR components). Nonetheless, conveying (parts of) the scenario using VR is likely to initiate the conversation, as the tool is deemed very desirable and participants are generally eager to try it out.

Additionally, VR is able to add a subjective user experience aspect to the conveyance of the potential plausible future of work scenario, allowing players to get immersed into story. Compared to using traditional methods for foresight (e.g., impression images and videos), more human-centered insights are likely to be gathered. Because of this, using VR adds unique value to the existing arsenal of tools for foresight.

For players to get immersed, the usability bottleneck needed to be overcome. Using the RITE framework during redevelopment, including specified observations of defined errors and failures, this job has been fulfilled satisfyingly. As per the requirements distilled from the literature study, which the VR experience satisfies, it can be reasonably assumed that the experience will be able to induce some level of presence (or immersion/flow). Specifically: players are likely to feel present in the mediated environment (the plausible future of work scenario), rather than in the immediate physical environment and have an optimal experience, being totally engaged, with the activity (of assembling bicycles).

As a degree of presence – immersion – flow is reached, natural behaviour and reactions, stimulated by the ‘real’ virtual environment, occurs. Having a representation of a plausible scenario allows much of the futures related uncertainties (concerning the scenario) to be researched. The potential for the use of VR is, when presence (and immersion/flow) is induced, not only limited to foresight for potential plausible futures, but also for futures studies in general, concerning potential possible futures (Dorsser et al., 2018).

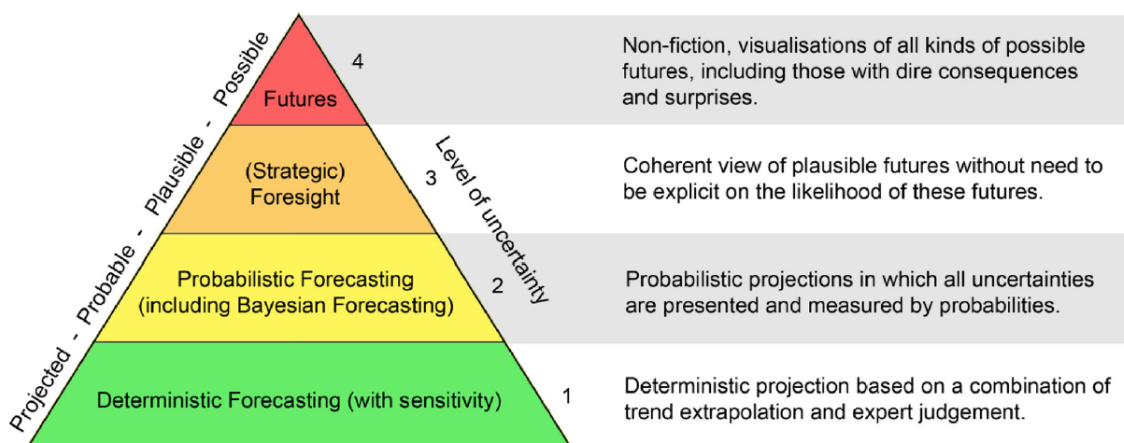


FIGURE 75. FUTURES PYRAMID (DORSSER ET AL., 2018)

Although it can be reasonably be assumed that the VR experience induces a degree of presence (or immersion/flow). A small test will be conducted using the presence questionnaire to evaluate this.

PRESENCE STUDY

STUDY DESIGN

In order to study the effect of the VR experience on presence (as defined by Steuer, 1995), the existing game will be used together with the redeveloped tutorial. As the requirements for immersion (distilled from the literature study) are fulfilled in version 1.6, evident from both the SUS scoring and the notable list of actions all participants were able to fulfil (without exception), the VR game should induce a measurable sense of presence (or immersion). Specifically, it is expected that the PQ score, obtained using the VR game and the revised version of the PQ (UQO Cyberpsychology Lab, 2004), will be comparable to the French norm added in the revision of the questionnaire. In other words, the null hypothesis is:

Ho: There is no difference in the mean values for presence between the norm and the VR experience, they are equal.

STIMULI

PQ in Qualtrics

The revised presence questionnaire (UQO Cyberpsychology Lab, 2004) created by Witmer and Singer (1994) will be used, so participants are able to self-report about the experience. The nineteen questions, divided over the five sub-scales *Realism*, *Possibility to Act*, *Quality of interface*, *Possibility to Examine* and *Self-evaluation of Performance*, were unaltered and used as is. The questionnaire has been copied into Qualtrics, where the requirements for this questionnaire were also added: participants are only able to answer the questions independently in the order they appear (i.e., questions cannot be skipped and participants cannot return to the previous question). A progress bar has additionally been added to convey the (remaining) length of the questionnaire to the participant. Screenshots of the questionnaire can be found in appendix F. Similar to the setup during RITE, a separate laptop will be used for the questionnaire.

VR game in Unity

Prior to entering the factory and the main sequence of the game, players will be eased into the game softly by learning the basic interaction controls and watching the clips concerning the story of the game (version 1.6 of the redeveloped tutorial). The flowchart for this part can be seen in figures 67 till 69. After, participants will assemble three 'half-bikes' (only one wheel is attached to the frame) in the factory. The flowchart of this part can be seen in figure 16 (starting from number 3). The experience will be played from Unity (version 2019.4.10f), using a separate laptop.

Procedure

Participants are welcomed and receive a brief introduction about the course of the study and into the VR game: "You are going to play a game where you'll assemble bicycles as a factory worker in a plausible future scenario. Think of it as playing in a world a couple of decades from now. Everything will become clear once you enter the game. This part of the study takes roughly 10 – 20 minutes. When finished, a questionnaire needs to be completed. This part takes roughly 10 minutes. Are you ready?". When the participant takes place in the middle of the play area, the researcher will help the participant put on the HMD and give the controllers. When ready, the game will be started on Unity, where players will go through the tutorial and the main sequence of the game. When finished, the participant will move from the VR play area to the second laptop to complete the PQ.

Participants

All participants are students and were recruited from the faculty of Industrial Design Engineering in Delft and from the surroundings of the Hague University. The education level ranged from vocational educational (MBO/HBO) to university (BSc/MSc) level. Eighteen participants were recruited in total.

DATA ANALYSIS

Data collected was recoded, changing the continuous scale from 1 – 7 to 0 – 6, and values concerning the quality of interface sub scale reversed (newValue = -1 x reversedItem + 6). Scores were consequently summed according to its sub scale, whereafter summary statistics like the mean (M) and standard deviation (SD) were calculated for each sub scale and the total score. The latter includes the mean and standard deviation of the sum of all sub scale scores.

In the revised PQ used (UQO Cyberpsychology Lab, 2004), a norm was included, containing the means and standard deviations of all sub scales and the total score. It was mentioned that these statistics were calculated using the data of 101 participants. A *t*-test will be conducted to compare the means (of the subscales and the total score) of the norm in the PQ and the VR experience. As it is assumed that the data is normally distributed, a significance level $\alpha = 0.05$ will be used.

Summary statistics are used for the test because the raw data of the norm in the PQ is not accessible. The *t*-test will be conducted through the syntax of SPSS, following the example provided by IBM on “how summary statistics can be used for a *t*-test”

(IBM, 2020; figure 76). Conducting the test through the GUI of SPSS is not possible as performing a *t*-test through the GUI requires the raw data of both datasets (the norm and the VR experience), so Levene’s test can simultaneously be conducted. Levene’s test examines whether homogeneity of variance can be assumed or not. This assumption is crucial as it determines which variant of the *t*-test (Welch’s test for unequal variances or *t*-test for equal variances) should be used. Because homogeneity of variance cannot be rejected nor accepted, both variants of the *t*-test will be conducted.

```

1 MATRIX DATA VARIABLES groups rowtype_ totalScore /factor=groups.
2 BEGIN DATA
3 1 N_SCALAR 101
4 2 N_SCALAR 18
5 1 MEAN 104,39
6 2 MEAN 87,39
7 1 SD 18,99
8 2 SD 10,36
9 END DATA.
10 ONEWAY
11 totalScore BY groups
12 /MATRIX=IN(*)
13 /CONTRAST=-1 1
14 /STATISTICS DESCRIPTIVES WELCH.
    
```

FIGURE 76. SPSS SYNTAX FOR T-TESTS

RESULTS

Summary statistics

Descriptives are calculated to see if the right value has been assigned to the right variable (M, SD and N per group) using the syntax. These can be seen in appendix G. Table 3 contains the values of the means (M), standard deviations (SD) and number of participants (N) of the norm and of the experience. With these values, the null hypotheses (mean values are equal and do not differ) is tested using Welch’s test first, followed by a *t*-test where homogeneity of variances is assumed.

	French Norm		VR Game	
	N	101	18	
	Mean	Std	Mean	Std
Realism	29.45	12.04	32.44	5.04
Possibility to act	20.76	6.01	19.28	2.30
Possibility to examine	15.38	4.90	13.89	2.83
Self-Evaluation of performance	11.00	2.87	10.22	1.31
Quality of interface	15.37	5.15	11.56	3.05
Total	104.39	18.99	87.39	10.36

TABLE 3. MEANS, STANDARD DEVIATIONS AND SAMPLE SIZES OF THE VR EXPERIENCE AND NORM

T-test results unequal variances assumed

With an α -level of 0.05,

- The mean value of presence (totalScore) of the norm (M = 104.39, SD = 18.99) seems to be **not** equal to the mean value of presence of the VR game (M = 87.39, SD = 10.36),
→ $t(40.958) = 30.315, p < .001$.
- The mean value of the realism sub scale of the norm (M = 29.45, SD = 12.04) seems to be equal to the mean value of the realism sub scale of the VR game (M = 32.44, SD = 5.04),
→ $t(58.821) = 3.141, p = .082$.
- The mean value of the possibility to act sub scale of the norm (M = 20.76, SD = 6.01) seems to be equal to the mean value of the possibility to act sub scale of the VR game (M = 19.28, SD = 2.30),
→ $t(66.745) = 3.362, p = .071$.
- The mean value of the possibility to examine sub scale of the norm (M = 15.38, SD = 4.90) seems to be equal to the mean value of the possibility to examine sub scale of the VR game (M = 13.89, SD = 2.83),
→ $t(38.166) = 3.252, p = .079$.
- The mean value of the self-evaluation of performance sub scale of the norm (M = 11.00, SD = 2.87) seems to be equal to the mean value of the self-evaluation of performance sub scale of the VR game (M = 10.22, SD = 1.31),
→ $t(52.104) = 3.442, p = .069$.
- The mean value of the quality of interface sub scale of the norm (M = 15.37, SD = 5.15) seems to be **not** equal to the mean value of the quality of interface sub scale of the VR game (M = 11.56, SD = 3.05),
→ $t(37.039) = 18.625, p < .001$.

Robust Tests of Equality of Means

totalScore				
	Statistic ^a	df1	df2	Sig.
Welch	30.315	1	40.958	.000

a. Asymptotically F distributed.

Realism				
	Statistic ^a	df1	df2	Sig.
Welch	3.141	1	58.821	.082

a. Asymptotically F distributed.

possibilityToAct				
	Statistic ^a	df1	df2	Sig.
Welch	3.362	1	66.745	.071

a. Asymptotically F distributed.

possibilityToExamine				
	Statistic ^a	df1	df2	Sig.
Welch	3.252	1	38.166	.079

a. Asymptotically F distributed.

selfEvalOfPerformance				
	Statistic ^a	df1	df2	Sig.
Welch	3.442	1	52.104	.069

a. Asymptotically F distributed.

qualityOfInterface				
	Statistic ^a	df1	df2	Sig.
Welch	18.625	1	37.039	.000

a. Asymptotically F distributed.

T-test results equal variances assumed

With an α -level of 0.05,

- The mean value of presence (totalScore) of the norm (M = 104.39, SD = 18.99) seems to be not equal to the mean value of presence of the VR game (M = 87.39, SD = 10.36),
→ $t(117) = 3.693, p < .001$.
- The mean value of the realism sub scale of the norm (M = 29.45, SD = 12.04) seems to be equal to the mean value of the realism sub scale of the VR game (M = 32.44, SD = 5.04),
→ $t(117) = 1.035, p = .303$.
- The mean value of the possibility to act sub scale of the norm (M = 20.76, SD = 6.01) seems to be equal to the mean value of the possibility to act sub scale of the VR game (M = 19.28, SD = 2.30),
→ $t(117) = 1.028, p = .306$.
- The mean value of the possibility to examine sub scale of the norm (M = 15.38, SD = 4.90) seems to be equal to the mean value of the possibility to examine sub scale of the VR game (M = 13.89, SD = 2.83),
→ $t(117) = 1.251, p = .214$.
- The mean value of the self-evaluation of performance sub scale of the norm (M = 11.00, SD = 2.87) seems to be equal to the mean value of the self-evaluation of performance sub scale of the VR game (M = 10.22, SD = 1.31),
→ $t(117) = 1.129, p = .261$.
- The mean value of the quality of interface sub scale of the norm (M = 15.37, SD = 5.15) seems to be not equal to the mean value of the quality of interface sub scale of the VR game (M = 11.56, SD = 3.05),
→ $t(117) = 3.038, p = .003$.

Contrast Tests^a

	Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
totalScore	Assume equal variances 1	-17.000000	4.6039066	-3.693	117	.000

a. Tests assuming unequal variances cannot be computed if matrix input is used.

	Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
Realism	Assume equal variances 1	2.990000	2.8899095	1.035	117	.303

a. Tests assuming unequal variances cannot be computed if matrix input is used.

	Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
possibilityToAct	Assume equal variances 1	-1.480000	1.4391243	-1.028	117	.306

a. Tests assuming unequal variances cannot be computed if matrix input is used.

	Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
possibilityToExamine	Assume equal variances 1	-1.490000	1.1913978	-1.251	117	.214

a. Tests assuming unequal variances cannot be computed if matrix input is used.

	Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
selfEvalOfPerformance	Assume equal variances 1	-.780000	.6907360	-1.129	117	.261

a. Tests assuming unequal variances cannot be computed if matrix input is used.

	Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
qualityOfInterface	Assume equal variances 1	-3.810000	1.2539121	-3.038	117	.003

a. Tests assuming unequal variances cannot be computed if matrix input is used.

CONCLUSION

Although the exact values of p differed between the two t -tests, with results from the t -test with equal variances assumed having generally more extreme values; conclusions drawn from both studies do not differ in terms of accepting or rejecting the null hypothesis. When it comes to presence (total score), Welch's variant of the t -test result (where homogeneity of variance was not assumed) is: $t(40.958) = 30.315$, $p < .001$. Result from the t -test variant where homogeneity of variance was assumed is: $t(117) = 3.693$, $p < .001$. The (null) hypothesis "*There is no difference in the mean values for presence between the norm and the VR experience, they are equal*" is therefore rejected in both test conclusions.

The same pattern arises when performing the same tests for the sub-scales of presence; although the p values from both variants of the t -test differ, conclusions drawn with both p values are the same (in terms of rejecting or accepting the null hypotheses). The sub scale quality of interface stands out during the t -tests, as it is the only scale (out of five) where the null hypothesis is rejected. Welch's test results for this sub scale is: $t(37.039) = 18.625$, $p < .001$. When homogeneity of variances is assumed, the result is: $t(117) = 3.038$, $p = .003$. The related (null) hypothesis "*There is no difference in the mean values for the quality of interface sub scale between the norm and the VR experience, they are equal*" is therefore rejected; there is a difference in the mean values for the quality of interface sub scale between the norms and the VR game.

Notably, the results of both variants of the t -test for the other four sub scales indicated that for the remaining four scales, the related null hypotheses are accepted; all having resulting p values > 0.05 .

- "*There is no difference in the mean values for the realism sub scale between the norm and the VR experience, they are equal*" is accepted.
- "*There is no difference in the mean values for the possibility to act sub scale between the norm and the VR experience, they are equal*" is accepted.
- "*There is no difference in the mean values for the possibility to examine sub scale between the norm and the VR experience, they are equal*" is accepted.
- "*There is no difference in the mean values for the self-evaluation of performance sub scale between the norm and the VR experience, they are equal*" is accepted.

The *quality of interface* sub scale is the only scale where the alternative hypothesis is accepted. Because the study was performed in a relatively small area, as described in the test setup (figure 43, redevelopment chapter), disruptive technology errors occurred more than a few times during the study (i.e., the need for a recharge of the controller, tracking issues caused by unfavourable placement of the lighthouses; causing lag and flicker issues and the cord connecting the HMD to the laptop getting in the way). The visual summary (figure 31), derived from the literature study, states that technology errors should be reduced as much as possible due to its considerable influence on the degree of presence (and immersion/flow) experienced. The (restricting) effect of technology errors is evident in this study; as the mean values of the total score (representing presence) for the norm and the VR experience are unequal due to significant differences in the mean values of the quality of interface sub scale.

The context of use is described in the context chapter. As the space in Robohouse is relatively more able to support the experience, having a bigger physical play area designated for VR, played on dedicated hardware (the VR experience can be played wirelessly using their VR tracker); it is assumed that similar problems (i.e., technology issues influencing the quality of interface negatively) will occur significantly less frequently (if at all). In other words, the quality of interface discrepancy measured is presumably caused by the (physical) test setup. Because the actual context of use is considerably bigger and also includes designated hardware, it is likely that the mean values will be equal if the test was performed at Robohouse.

With this assumption, every sub scale that makes up the total score for presence in the VR experience would have means that are equal to the norm. As the score for presence is calculated through summing up the sub scale scores, the mean value for presence in the VR experience would consequently (presumably) show no significant differences with the norm; both having equal means, indicating a similar experience of presence.

LIMITATIONS

Using the *t*-test requires several assumptions to be true. The data in each group is obtained through random sampling. Furthermore, the data collected is continuous (or ordinal) and is normally distributed for both groups. The sample size used is 'adequate' and the homogeneity of variances is true (when using the variant where homogeneity of variance is assumed).

Because both variants of the *t*-test are used, when homogeneity of variances can be assumed and when homogeneity of variances cannot be assumed, the assumption of homogeneity of variance is less important. Furthermore, as the PQ only includes 7-point Likert scales, data collected is continuous and likely normally distributed.

However, as the sample size used for this study was 18 (participants), compared to the 101 participants used during the revision of the PQ (creating the norm), the assumption that the sample size is adequate might not hold up. Furthermore, in the conclusion, it is assumed that the quality of interface issues experienced during this study, will not occur when the VR experience is played using the dedicated hardware and VR play area of Robohouse. The rejection of the null hypothesis (there is no difference in mean values for the quality of interface sub scale) is argued to be caused solely by the test setup. As this was not tested, the assumption might also not hold up.

Because of this, the conclusions are of little statistical relevance and result should be only regarded as preliminary. Nonetheless, as the VR experience satisfies all criteria for immersion to take place (distilled from the literature study), no failures and errors were observed using the latest version of the tutorial (version 1.6) and preliminary results indicate that players are likely to experience presence (to a similar degree as the norm), there is little reason to assume otherwise.

VIRTUAL REALITY FOR RESEARCH

A digital twin is a virtual representation that serves as a digital copy of a physical object or process. Because the VR experience serves as a digital representation of a potential plausible (or possible) scenario, there is no physical object or process (yet), as the scenario takes place in the future. Nonetheless, as the VR experience is able to immerse the players into the future, it becomes more feasible and valuable to be used in futures studies for potential possible futures as well. For instance, through virtual prototyping, a significant amount of resources used (related to physical prototyping, like materials), can be reduced. Furthermore, virtual representations can help support researchers in exploring potential risky situations and (human-robot) interactions. Additionally, ergonomic considerations (like vulnerabilities or fittings) can be uncovered and taken into account more easily.

VERDICT

The power of a human centered VR experience of the future of work is its ability to immerse players into the content. This makes it a valuable foresight tool for (plausible) futures, as players are able to experience the content (as opposed to view the content), but also a valuable tool for futures studies concerning possible futures. As a degree of presence – immersion – flow is reached, VR can help assist researchers study more dangerous situations and take in account ergonomic considerations much more easily. Because the representation (of the future) is virtual, this will also amount to much resources related to physical prototyping saved.

Futures are hard to imagine and often difficult to study cohesively and consistently. The immersive potential of VR is able to bridge these difficulties while also adding a subjective component to what otherwise would've been more objective. Thus, VR is able to become a powerful tool for futures studies, in order to help us **design for our future**.



References

- Admiraal, W., Huizenga, J., Akkerman, S., & Dam, G. T. (2011). The concept of flow in collaborative game-based learning. Retrieved from: doi:10.1016/j.chb.2010.12.013
- Arns, L. L., & Cerney, M. M. (2005). The Relationship Between Age and Incidence of Cybersickness Among Immersive Environment Users.
- Biocca, F., & Delaney, B. (1995). Immersive virtual reality technology. Communication in the age of virtual reality. Lawrence Erlbaum Associates, Inc. (Reprinted in modified form from "Journal of Communication," Aut 1992)
- Brook, J. (1996). SUS: A Quick and Dirty Usability Scale. (1996). Retrieved from: doi:10.1201/9781498710411-35
- Chavez B., Bayona S. (2018) Virtual Reality in the Learning Process. In: Rocha Á., Adeli H., Reis L., Costanzo S. (eds) Trends and Advances in Information Systems and Technologies. WorldCIST'18 2018. Advances in Intelligent Systems and Computing, vol 746. Springer, Cham. https://doi.org/10.1007/978-3-319-77712-2_129
- Cowley, B., Charles, D., Black, M., & Hickey, R. (2008). Toward an understanding of flow in video games. Retrieved from: doi:10.1145/1371216.1371223
- Csikszentmihalyi, M. (1991). Flow: The Psychology of Optimal Experience. Retrieved from: doi:10.2307/258925
- Csikszentmihalyi, M. (1997). Finding Flow: The Psychology of Engagement with Everyday Life.
- van Dorsser, C., Walker, W. E., Taneja, P., & Marchau, V. A. (2018). Improving the link between the futures field and policymaking. *Futures*, 104, 75–84. <https://doi.org/10.1016/j.futures.2018.05.004>
- Fang, X., Zhang, J., & Chan, S. S. (2013). Development of an Instrument for Studying Flow in Computer Game Play. Retrieved from: doi:10.1080/10447318.2012.715991
- ISO 9241-210. (2019). Retrieved from <https://www.iso.org/standard/77520.html>
- Jin, S. A. (2011). I Feel Present. Therefore, I Experience Flow: A Structural Equation Modeling Approach to Flow and Presence in Video Games. Retrieved from: doi:10.1080/08838151.2011.546248
- Joshi, S., Hamilton, M., Warren, R., Faucett, D., Tian, W., Wang, Y., & Ma, J. (2021). Implementing Virtual Reality technology for safety training in the precast/ prestressed concrete industry. *Applied Ergonomics*, 90, 103286.
- Kaur, K. (1997). Designing Virtual Environments for Usability. *Human-Computer Interaction*. Retrieved from: https://link.springer.com/content/pdf/10.1007/978-0-387-35175-9_112.pdf
- Kiili, K. (2005). Content creation challenges and flow experience in educational games: The IT-Emperor case. Retrieved from: doi:10.1016/j.iheduc.2005.06.001
- Kim, Y. Y., Oh, S., & Lee, H. (2005). What makes people experience flow? Social characteristics of online games. Retrieved from: doi:10.1504/ijamc.2005.007724
- Liu, C., & Uang, S. (2007). Measurement and Prediction of Cybersickness on Older Users Caused by Virtual Environments.
- Liu, D., Wang, T. (2006). A Virtual Reality Training System for Robot Assisted Neurosurgery. *Advances in Artificial Reality and Tele-Existence*.
- Medlock, M. C., Wixon, D., Terrano, M., Romero, R. L., Fulton, B. (2007). Using the RITE method to improve products; a definition and a case study. Retrieved from: <https://www.semanticscholar.org/paper/Using-the-RITE-method-to-improve-products%3B-a-and-a-Medlock-Wixon/5340ef8a91900840263a4036b0433a389b7097b2>
- Mousavi, M., Jen, Y. H., & Musa, S. N. B. (2013). A Review on Cybersickness and Usability in Virtual Environments. *Advanced Engineering Forum*, 10,

- 34–39. <https://doi.org/10.4028/www.scientific.net/aef.10.34>
- Mütterlein, J. (2018). The Three Pillars of Virtual Reality? Investigating the Roles of Immersion, Presence, and Interactivity. Retrieved from: <https://core.ac.uk/download/pdf/143481013.pdf>
 - Nielsen, J. (2012). How Many Test Users in a Usability Study? Nielsen Norman Group. <https://www.nngroup.com/articles/how-many-test-users/>
 - Nah, F. F. H., Eschenbrenner, B., Zeng, Q., Telaprolu, V. R., & Sepehr, S. (2014). Flow in gaming: literature synthesis and framework development. *International Journal of Information Systems and Management*. Retrieved from: <https://doi.org/10.1504/ijisam.2014.062288>
 - R. S. Kennedy, R. S., Lane, N. E., Berbaum, K. S., Lilienthal, M.G. (1993). Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. Retrieved from: [doi: 10.1207/s15327108ijap0303](https://doi.org/10.1207/s15327108ijap0303)
 - Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgemant, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*. Retrieved from: <https://doi.org/10.1016/j.compedu.2019.103778>
 - Riccio, G. E., & Stoffregen, T. A. (1991). An ecological Theory of Motion Sickness and Postural Instability. Retrieved from: https://doi.org/10.1207/s15326969eco0303_2Instability
 - RoboValley. (2020, May 17). About. Retrieved from: <https://robovalley.com/connect/about/>
 - RoboValley. (2021a). Rapid Robotics for Operator Safety: What a Bottle Picker Can Do. Retrieved from: <https://robovalley.com/activities/news/what-a-bottle-picker-can-do/>
 - RoboValley. (2021b). Demo Day: the most exciting day of the year. Retrieved from: <https://robovalley.com/activities/news/demo-day-the-most-exciting-day-of-the-year/>
 - Ryan, M.L. (2015). Narrative of Virtual Reality. Retrieved from: <http://www.movingimages.info/class/wp-content/uploads/2010/06/RyNarr.pdf>
 - Schmitt, J., Hillenbrand, A., Kranz, P., Kaupp, T. (2021). Assisted Human-Robot-Interaction for Industrial Assembly. *HRI '21 Companion*. Retrieved from: <https://doi.org/10.1145/3434074.3447127>
 - Slater, M., & Wilbur, S. (1997). A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. Retrieved from: [doi:10.1162/pres.1997.6.6.603](https://doi.org/10.1162/pres.1997.6.6.603)
 - Smith, S., Marsh, T., Duke, D., Wright, P. (2001). Drowning in immersion. Retrieved from: https://www.researchgate.net/publication/2362590_Drowning_in_Immersion
 - Stanney, K. (1995). Realizing the full potential of virtual reality: human factors issues that could stand in the way.
 - Steuer, J. (1995). Defining Virtual Reality: Dimensions Determining Telepresence. Retrieved from: <https://steinhardtapps.es.its.nyu.edu/create/courses/2015/reading/steuer.pdf>
 - Sweetser, P., & Johnson, D. (2004). Player-Centered Game Environments: Assessing Player Opinions, Experiences, and Issues. Retrieved from: [doi:10.1007/978-3-540-28643-1_40](https://doi.org/10.1007/978-3-540-28643-1_40)
 - Sweetser, P., & Wyeth, P. (2005). GameFlow: a model for evaluating player enjoyment in games. Retrieved from: [doi:10.1145/1077246.1077253](https://doi.org/10.1145/1077246.1077253)
 - Technische Universiteit Delft. (2021, February 2). Making a search plan. TUlib. Retrieved from: <https://tulib.tudelft.nl/searching-resources/making-a-search-plan/>
 - IBM. (2020). T-Test with summary statistics in SPSS. IBM Support. <https://www.ibm.com/support/pages/t-test-summary-statistics-spss>
 - Turner, C.W., & Lewis, J.R., & Nielsen, J. (2006). Determining Usability Test Sample Size. *International Encyclopedia of Ergonomics and Human Factors*, 2006, Second Edition, Volume 3
 - Voros, J. (2003). A generic foresight process framework. *Foresight*, 5(3), 10–21. <https://doi.org/10.1108/14636680310698379>
 - Voros, J. (2017). The Futures Cone, use and history. Blog: The voroscope, <https://thevoroscope.com/2017/02/24/the-futures-cone-use-and-history/>.
 - Walker, W. E., Marchau, V. A. W. J., & Kwakkel, J. H. (2013). Uncertainty in the framework of policy analysis. In W. A. H. Thissen, & W. E. Walker (Eds.). *Public policy analysis: New developments*. New York: Springer.
 - Wann, J., Williams, M. M. (1996). What does virtual reality need? Human factors issues in the design of three-dimensional computer environments. Retrieved from: <https://www.sciencedirect.com/science/article/abs/pii/S107158199690035X>
 - Witmer, B.G., & Singer, M.J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire.

