Master Thesis in

Human Computer Interaction

Diverse Perspectives: An Inclusive Serious Game for International and Dutch Students at Utrecht University

Author:
Geertje Peters Rit
4511298
g.w.f.x.petersrit@students.uu.nl

First supervisor/assessor:
Dr. Hanna Hauptmann
Department of Information and Computing Sciences
h.j.hauptmann@uu.nl

Second supervisor/assessor:
Prof. dr. Anja Volk
Department of Information and Computing Sciences
a.volk@uu.nl

Utrecht University

Department of Information and Computing Sciences

August 2, 2023

Abstract

Nowadays, higher education has a diverse student population. Increased globalisation and more educational and financial opportunities are among the reasons why. Although there is a greater variety of students' background, skills and abilities, this does not necessarily mean they have more inclusive interactions. In other words, while the diversity in the student population increased, people do not seem to use it to their benefit. Therefore, how could a serious game be used to promote diversity at Utrecht University (UU)?

Attempting to answer the research question, this thesis created an inclusive serious game, called *Diverse Perspectives*. The game is a visual novel with multiple endings where the player's choices affect the outcome. The narrative of *Diverse Perspectives* follows a student on their very first day at UU. It contains several dilemmas in the context of diversity and (international) student inclusion. *Diverse Perspectives* aims to raise awareness about the many aspects of diversity. Additionally, it allows the player to think and reflect on them, as the intended message of the game is to promote diversity.

Designing *Diverse Perspectives* followed an iterative and user-centred approach. A low-fidelity prototype was made and evaluated. The target users (UU students) were involved early in the design process. During the experiment, the high-fidelity prototype of *Diverse Perspectives* was used. This prototype consists of two versions: *with* the perspective switch (version A) and *without* the switch (version B). Furthermore, a mixed-methods approach with a between-subjects design was chosen to collect data from 32 participants. The qualitative data consists of semi-structured interviews, while the quantitative data contains the results of the Shortened Version of the Miville-Guzman Universality-Diversity Scale (M-GUDS-S) and Game Experience Questionnaire (GEQ). These measure the player's attitude towards diversity, and gaming experience, respectively. The experiment had two parts, where part one was in-person. Part two followed a week later and was completed online.

Moreover, we hypothesised a positive change in the M-GUDS-S after playing the game and that this difference would be bigger in version A. Likewise, we hypothesised that version A would score higher on the GEQ than version B. Lastly, we hypothesised that players who identify with a marginalised group would score higher on the GEQ as well, compared to players who do not.

Although we obtained three significant results, our quantitative data could not confirm our hypotheses. Nevertheless, the qualitative data revealed insightful information on the participants' experience of the game. The perspective switch was perceived as helpful, since it created understanding for the point of view of others. Likewise, eight preliminary theories have been formed that seem promising to study in the future, as well as the results from the exploratory research. Moreover, a lot of suggestions have been offered to improve *Diverse Perspectives*. While participants enjoyed the game and perceived it as educational, it has many opportunities for improvement.

Acknowledgements

I would like to thank some people that supported me during the whole thesis process. First, my supervisor Dr. Hanna Hauptmann for giving me the freedom to shape my own research, the feedback at the weekly meetings and connecting me with my second supervisor and two other students. I really appreciated your active involvement and encouragement.

Next, I would like to thank my second supervisor Prof. dr. Anja Volk for the feedback at the joint meetings in the second phase of the process, making me aware of diversity-related events and the general enthusiasm.

A big thanks to Sinie van der Ben and Marc Ferriggi. I really liked being able to share my worries about all things, big and small, or just having a chat about anything. It was nice to get to know you both.

A huge thank you to my bi-weekly study session group: Chandni Bagchi, Thirza Hiwat and Paulina Moerland. You helped me stay motivated and the snacks did their part as well.

Furthermore, I would like to thank my *Serious Gaming* group for letting me continue our project from the course last year. I intended to not make another *Honeymoon*, I promise.

Additionally, I would like to thank my parents for the emotional support and having my back, so that I could focus on my thesis.

Last, but not least, I would like to thank all the participants for their participation, with a special thanks to Tamara Mantz.

Contents

List of Figures **vii**

List of Tables **x**

List of Acronyms **xii**

1 Introduction **1**
1.1 Research problem and motivation 2
1.1.1 Research goal 2
1.1.2 Research questions 3
1.2 Contribution 4
1.3 Outline 4

2 Related work **6**
2.1 What are serious games? 7
2.1.1 Difference between entertainment games, serious games, gamification and game-based learning 7
2.1.2 Different domains of serious games 7
2.1.3 The theory behind serious games 9
2.1.4 Pedagogical learning theories 11
2.2 Storytelling 13
2.2.1 Different types of narratives and progression 13
2.2.2 Game design principles for visual novels 15
2.3 Serious games in the domain of inclusion and diversity 16
2.3.1 Change Perspective 16
2.3.2 MASELTOV 18
2.3.3 The Dilemma Game 19
2.3.4 Other related serious games 20
2.4 The role of diversity 21
2.4.1 What is diversity? 21
2.4.2 How to encourage diversity? 23
2.5 Main findings of the Related Work 24

3 Design **25**
3.1 Game mechanics 25
3.2 Content of the story 26
3.2.1 Narrative structure 26

3.2.2 Take-away messages at the end of the game . . . 27
3.3 High level overview of the iterative design procedure . . . 28
3.3.1 Initial interviews . . . 28

4 Implementation **31**
4.1 Low-fidelity prototype . . . 31
4.1.1 Evaluation of the low-fidelity prototype . . . 32
4.2 High-fidelity prototype . . . 33
4.2.1 Story . . . 33
4.2.2 Endings . . . 34
4.2.3 Choices . . . 34
4.2.4 Avatar sprites and behaviour . . . 36

5 Method **38**
5.1 Participants . . . 38
5.2 Materials and measurements . . . 39
5.2.1 Independent variables . . . 41
5.2.2 Dependent variables . . . 41
5.2.3 Controlling variables . . . 41
5.3 Procedure . . . 41

6 Results **43**
6.1 Quantitative analysis . . . 43
6.1.1 Hypotheses . . . 43
6.1.2 Exploratory research . . . 50
6.1.3 Choice distribution flow . . . 56
6.2 Main findings of the quantitative analysis . . . 59
6.3 Qualitative analysis . . . 59
6.3.1 Open and axial coding . . . 59
6.3.2 Selective coding . . . 64
6.4 Main findings of the qualitative analysis . . . 68

7 Discussion **69**
7.1 Interpretation of the results . . . 69
7.1.1 Quantitative results . . . 69
7.1.2 Qualitative results . . . 71
7.2 Limitations . . . 74
7.2.1 Participant sample . . . 74
7.2.2 Risk of socially desired answers . . . 74
7.2.3 Risk of researcher's bias . . . 75
7.2.4 The follow-up survey . . . 75
7.3 Further research . . . 76
7.3.1 Actual distribution of the game . . . 76
7.3.2 Game improvements . . . 77
7.3.3 Player diversity . . . 78
7.4 Main findings of the Discussion . . . 78

8 Conclusion **80**

A Figures **87**
A.1 Narrative branching structure . . . 87
A.2 Sankey flow chart . . . 89
A.3 Hierarchical code structure . . . 91

B Tables **97**
B.1 Descriptive statistics M-GUDS-S . . . 97
B.2 Descriptive statistics GEQ . . . 99
B.3 Descriptive statistics M-GUDS-S exploratory . . . 100
B.4 Descriptive statistics GEQ exploratory . . . 101

C Initial interviews **103**
C.1 Materials . . . 103
C.2 Semi-structured Interview . . . 103
C.2.1 Opening the interview . . . 103
C.2.2 Introducing the research . . . 103
C.2.3 Beginning the interview . . . 103
C.2.4 Main part of the interview . . . 104
C.2.5 Closing the interview . . . 105

D Focus Groups **106**
D.1 Demographic questionnaire . . . 106
D.2 Topic list . . . 107

E Experiment **109**
E.1 Materials . . . 109
E.2 Semi-structured Interview . . . 109
E.2.1 Opening the interview . . . 109
E.2.2 Introducing the research . . . 109
E.2.3 Beginning the interview . . . 109
E.2.4 Main part of the interview . . . 110
E.2.5 Closing the interview . . . 111

F Shortened Version of the Miville-Guzman Universality-Diversity Scale (M-GUDS-S) questionnaire **112**

G Game Experience Questionnaire (GEQ) **114**

H Information Sheet and Informed Consent Form **116**
H.1 Information Sheet . . . 116
H.1.1 Study Information . . . 116
H.1.2 Purpose [initial interviews] . . . 116
H.1.3 Purpose [focus groups] . . . 116
H.1.4 Purpose [experiment] . . . 116
H.1.5 Your Rights . . . 116
H.2 Consent Form . . . 117

I Storyline **118**
I.1 Scene: At home 118
I.2 Travelling 119
I.3 At the UU campus 122
I.4 In the lecture room 127
I.5 In KBG 131
I.6 In the city centre 133
I.7 In a supermarket 137
I.8 At the *hospi* 139
I.9 Recap 139

List of Figures

2.1 An overview of the classification of serious games distributed along five categories, created by Laamarti et al. (2014). A game could consist of multiple sub-categories per category. 8

2.2 The dual flow in serious games, created by Dörner et al. (2016). The original figure contained a typo ('maldadaption'), which has been corrected here. When the task difficulty is too high, players move in the anxiety zone. When the task difficulty is too low players move in the boredom zone. When the task difficulty and the skill level are in balance, the player is in the flow. This flow can vary in degrees of low, medium and high. 11

2.3 Screenshots of *Change Perspective* from the perspective of the German student Lena (a) and the Dutch student Kim (b). The player interacts with the game via a point-and-click system (Backx et al., 2022). 17

2.4 A dilemma from the *Dilemma Game* (Equality, Diversity & Inclusion (EDI) Office of Utrecht University, 2022). This dilemma falls under the 'Identity' theme and the 'Sexual orientation' subtheme. Each dilemma has four options and players should share with each other which option they favour and why. . 20

2.5 The Academic Wheel of Privilege. It shows a non-exhaustive overview of diversity presented in seven different categories that are divided into twenty aspects. The closer an individual is located to the inner ring, the more likely they can benefit from privileges (Elsherif et al., 2022). 22

4.1 Screenshots of Dilemma #3A in *Diverse Perspectives*. The player encounters a situation (a) and is then offered to choose a response option (b). 35

4.2 Screenshots of trivial choices in *Diverse Perspectives*. The player can choose what to do after their alarm went off (a) and what drink they would like in the *Spar University* (b). 36

6.1 The distributions of the pre-test (blue) and post-test (yellow) M-GUDS-S data in a box plot. The three components (Diversity of Contact (DC), Relativistic Appreciation (RA) and Comfort with Differences (CD)) are on the x-axis. The score of each component is on the y-axis. 45

6.2 The distributions of the score differences of the M-GUDS-S in a box plot between version A (blue) and version B (yellow). The three components (DC, RA and CD) are on the x-axis. The difference score of each component is on the y-axis. 46

6.3 The distributions of the GEQ in a box plot between version A (green) and version B (purple). The four components (Immersion, Flow, Competence and Positive Affect) are on the x-axis. The score is on the y-axis. 47
6.4 the distributions of the GEQ in a box plot per group, so identification with a marginalised group (Y, green) and no identification with a marginalised group (N, purple). The four components (Immersion, Flow, Competence and Positive Affect) are on the x-axis. The score is on the y-axis. 49
6.5 The distributions of the score differences of the M-GUDS-S in a box plot per group, so identification with a marginalised group (Y, blue) and no identification with a marginalised group (N, yellow).The three components (DC, RA and CD) are on the x-axis. The difference score of each component is on the y-axis. 51
6.6 The distributions of the score differences of the M-GUDS-S in a box plot between Females (blue) and Males (yellow). The three components (DC, RA and CD) are on the x-axis. The difference score of each component is on the y-axis. 52
6.7 The distributions of the GEQ in a box plot between Females (green) and Males (purple). The four components (Immersion, Flow, Competence and Positive Affect) are on the x-axis. The score is on the y-axis. 53
6.8 The distributions of the score differences of the M-GUDS-S in a box plot between Dutch (blue) and Abroad (yellow). The three components (DC, RA and CD) are on the x-axis. The difference score of each component is on the y-axis. 54
6.9 The distributions of the GEQ in a box plot between Dutch (green) and Abroad (purple). The four components (Immersion, Flow, Competence and Positive Affect) are on the x-axis. The score is on the y-axis. 55
6.10 The Sankey flow chart of the choices players made and the dilemmas, with a hashtag (#), they encountered. The number after each label represents the number of players that encountered, where 32 is the maximum number. The width of a branch indicates the number of people that made a certain choice. 57
6.11 The main categories after open and axial coding. The circled number underneath each parent represents the number of child codes. 60
6.12 The five tree diagrams of the hierarchical code structure. The number at the bottom of the child nodes indicates the number of participants that mentioned that specific code. 61

A.1 Snippets of the narrative branches as made in Miro. The arrows go from left to right. Consequential choices are represented by green speech bubbles, while trivial choices are displayed as red diamonds. Orange rectangles show choice options and blue stars denote a perspective switch. Lastly, black boxes depict dialogue scenes. 88
A.2 The Sankey flow chart, which has been made in SankeyMATIC. It shows the choices players made and the dilemmas, with a hashtag (#), they encountered. The number after each label represents the number of players that encountered, where 32 is the maximum number. The width of a branch indicates the number of people that made a certain choice. 90

A.3 All the children codes of the five parent codes. The number at the bottom of the child nodes indicates the number of participants that mentioned that specific code. The tree diagrams of the hierarchical code structure are made in Xmind. 96

List of Tables

3.1 The steps of the iterative design for *Diverse Perspectives*. Each step is accompanied with its purpose, participant pool and aimed number of participants. . 28

5.1 The demographics of the participants (N = 32) who participated in the experiment. 38
5.2 Timeline of part 1 of the experiment. This part was done in-person. The total estimated time of the timeline is 45 min. 42
5.3 Timeline of part 2 of the experiment. This part was done online. The total estimated time of the timeline is 5-10 min. 42

6.1 Overview of the analyses that have been conducted per questionnaire. 44
6.2 The player statistics regarding the achieved endings. 57
6.3 The percentage distributions of choices made in each dilemma. The choice options are ordered 'Good-Neutral-Bad'. 58

B.1 The descriptive statistics representing the scores for the M-GUDS-S pre-test and post-test data along the three components for the whole sample. 97
B.2 The descriptive statistics representing the scores for the M-GUDS-S pre-test and post-test data along the three components for version A (the experimental condition). 98
B.3 The descriptive statistics representing the scores for the M-GUDS-S pre-test and post-test data along the three components for version B (the control condition). 98
B.4 The descriptive statistics representing the scores for the M-GUDS-S pre-test and post-test data along the three components for the two conditions. 99
B.5 The descriptive statistics representing the scores for the GEQ data along the four components for both conditions. 99
B.6 The descriptive statistics representing the scores for the GEQ data along the four components for participants who identify with a marginalised group (Y) and participants who do not (N). 100
B.7 The descriptive statistics representing the scores for the M-GUDS-S pre-test and post-test data along the three components for participants who identify with a marginalised group (Y) and participants who do not (N). 100
B.8 The descriptive statistics representing the scores for the M-GUDS-S pre-test and post-test data along the three components for the female group (F) and the male group (M). 101

B.9 The descriptive statistics representing the scores for the M-GUDS-S pre-test and post-test data along the three components for the Dutch group (A) and the abroad group (A). 101
B.10 The descriptive statistics representing the scores for the GEQ data along the four components for the female group (F) and the male group (M). 102
B.11 The descriptive statistics representing the scores for the GEQ data along the four components for the Dutch group (D) and the Abroad group (A). 102

D.1 Timeline of the focus group session, which should be done in-person. The total estimated time of the timeline is 60 min. 106

List of Acronyms

1PP First-Person Perspective

3PP Third-Person Perspective

8LEM Eight Learning Events Model

AI Artificial Intelligence

ARCS Attention, Relevance, Confidence and Satisfaction

CD Comfort with Differences

DC Diversity of Contact

ECL Extraneous Cognitive Load

GCL Germane Cognitive Load

GEQ Game Experience Questionnaire

HCI Human-Computer Interaction

ICL Intrinsic Cognitive Load

MAPP MASELTOV App

MASELTOV Mobile Assistance for Social Inclusion and Empowerment of Immigrants with Persuasive Learning Technologies and Social Network Services

M-GUDS-S Shortened Version of the Miville-Guzman Universality-Diversity Scale

NPC Non-Player Character

PENS Player Experience of Need Satisfaction

RA Relativistic Appreciation

SDT Self-Determination Theory

UU Utrecht University

UX User Experience

Chapter 1

Introduction

Nowadays, higher education has a diverse student population. Increased globalisation and more educational and financial opportunities are among the reasons why (Homan, 2019; Trolian & Parker III, 2022). Since humans are social beings, they feel an inherent need to belong and thus form groups (Allen et al., 2022; Baumeister & Leary, 1995). A group is a collection of two or more people that see themselves as a member of the same association. When people form groups, they tend to include persons that are similar to them. In other words, this implies that people have the tendency to exclude others that do not have much in common with them. These observations are in line with the *Similarity/Attraction Theory* (Byrne, 1997).

According to the *Social Identity Theory* from Tajfel and Turner (1986), people first align with a group based on categorisation. Once they have categorised themselves in a group, they identify with the group's norms and values. Then, they compare their 'inside' group with other groups, the so-called 'outside' groups. The construction of these inside and outside groups could create a 'we-vs-them' feeling, where people develop an intergroup bias. A bias is a prejudice for or against something, based on (in)visible characteristics. This bias is positive for people who belong to the same group, in other words, the ingroup. However, the other side of the coin is that people hold a negative bias against others who are not a member of this inside group, that is, the outgroup.

Thus, although there is a greater variety of students' background, skills and abilities, this does not necessarily mean students have more inclusive interactions. This could be caused due to the outgroup bias, as mentioned before. These negative biases could contribute to inequities in many different domains, such as education, health and employment. Consequently, these barriers could cause negative health consequences, such as stress (Dovidio et al., 2017; Kukulska-Hulme et al., 2015). Thus, while the diversity in the student population increased, people do not seem to use it to their benefit.

Even more, making people aware of diversity could have the opposite effect. Gómez-Zará et al. (2020) investigated team formation under the influence of a 'diversity score'. This score would denote how 'diverse' a team was, based on the (dis)similarities between them. The diversity score would increase when people added others to their team they had little in common with. Likewise, the score would decrease when people teamed up with others that were similar to them. Their results indicated when this diversity score was visible, people were less likely to collaborate with others that differed from them. In other words, the score even increased their preference for similarity.

1.1 Research problem and motivation

Hence, the problem is that there is little interaction between people that differ from one another. Consequently, groups are not as diverse as they could be. People prefer to stay in their own social bubbles. When there is little to no diversity, a serious consequence could be the underrepresentation of marginalised groups (Puritty et al., 2017). These are groups that are treated as less important based on their (in)visible characteristics. Underrepresentation could induce severe repercussions, such as negative stereotypes, discrimination and lack of inclusion (Moreu et al., 2021).

Zooming in on a specific scope, consider the different groups that students form at any Dutch university. Generally, internationals stick together and the same holds for Dutch students (Rienties & Tempelaar, 2013). According to *Contact Theory*, interacting with others who are different from one causes greater openness to diversity and vice versa (Dovidio et al., 2017). Diversity is beneficial for teamwork and less automatic categorisation of groups, thus promoting it would be favourable (Homan, 2019). Moreover, by being aware how others are similar, yet still differ from one is crucial for effective interaction with them. That is, one could form alliances with others who are similar, while simultaneously being able to accept and value people that differ from oneself (Fuertes et al., 2000). In other words, by recognising similarities, it allows one to feel connected with others, while one still sees oneself as a unique individual, because everyone is different.

Hence, how could the formation of diverse groups be encouraged? An answer could be through the means of serious games, as they allow the player to see things from another perspective. This could invoke empathy and understanding for people in a different group (Belman & Flanagan, 2010) and hence stimulate intergroup interactions.

To further narrow down the scope, the focus of this thesis will be centred on Utrecht University (UU) students. This specific scope is chosen, because this thesis continues the work of *Change Perspective*. *Change Perspective* is a narrative-based serious game targeted at international and Dutch students. The game has been made in the *Serious Gaming (INFOMSEGA)* course in 2022 at UU (Backx et al., 2022). The reason why it focused on international and Dutch student groups is that there were two international students in the project group. They remarked they felt a bit like an outsider at times, especially in an all-Dutch group. This issue does not only play a role in Utrecht, but seems to apply to a wider scope as well (Rienties et al., 2014; Rienties & Tempelaar, 2013).

To face this problem, one of the initial ideas was to make a social game that would teach internationals about Dutch culture. Nevertheless, this would place the burden of education again on the international students, which was not desirable. Furthermore, if the game was targeted solely at internationals, it would make it less interesting to play for Dutch students. Hence, *Change Perspective* shows two different perspectives, namely that of the international student and of the Dutch student. Moreover, *Change Perspective* has been inspired by the *Dilemma Game* (Equality, Diversity & Inclusion (EDI) Office of Utrecht University, 2022). Section 2.3.1 will elaborate more on *Change Perspective*, while section 2.3.3 will discuss the *Dilemma Game* in greater detail.

1.1.1 Research goal

The research goal of the thesis is to promote diversity at UU through serious games, with the vision of mitigating negative intergroup biases. It aims to do so by extending *Change*

Perspective so that it becomes an inclusive game. Although this term is usually associated with accessibility, here 'inclusive' means the representation of diverse people. In other words, anybody should feel included when playing the game. The terms 'diversity' and 'inclusion' are related, although they are not the same. While the first term denotes the (in)visible characteristics that make people different from each other, the second term stands for creating a place where each person feels welcomed and valued, regardless of their differences. In other words, it is still possible to have diversity without any inclusion. Diversity could rather be seen as a stepping stone to inclusion. Inclusion consists of two parallel needs: uniqueness and belonging. It is not easy to meet both needs, as the feeling of belonging could require a person to sacrifice a bit of their uniqueness to blend in (Serebrenik, 2022).

The ultimate goal of the game is to encourage inclusive behaviour. That is, the intermingling and interaction between different student groups in a manner that appreciates diversity. While diversity consists of observable aspects, inclusion is based on a subjective feeling. Since inclusion is more difficult to measure, we chose to measure diversity instead. Hence, the focus will be more on diversity. Nevertheless, the game is still intended to be an inclusive game. Diversity in this context means representing the (in)visible differences between people, without implying any vertical ranking or reinforcing stigmatising stereotypes. Rather, differences between people should be seen as complementary to conquer the bias against diversity (Braidotti, 2011). According to *Information Decision-Making Theory*, diversity may benefit the amount of knowledge and skills of the group (Gómez-Zará et al., 2020; Pitz & Harren, 1980). *Diverse Perspectives*, the upgraded version of *Change Perspective*, could contribute to teaching these lessons.

1.1.2 Research questions

To satisfy the research goal, the following research question has been formulated: *How could Diverse Perspectives be used to promote diversity at Utrecht University?*

The research question is divided into three parts. Together, these parts form an answer to the research question. These subquestions are:

- *R1: What is the impact of the perspective switch on one's attitude towards diversity?*
- *R2: Which factors contribute to a positive gaming experience?*
- *R3: What is the effect of identifying with a marginalised group on the gaming experience?*

The motivation for choosing a serious game as a format is because they provide a platform for incidental and observational learning at the user's own pace, so they stay motivated enough to continue playing while receiving the informational content. Incidental learning is "unintentional or unplanned learning that results from other activities" (Kukulska-Hulme et al., 2015, p.10) and occurs during daily tasks. It is often triggered by context. Knowledge obtained from this type of learning is perceived as beneficial for self-confidence and self-knowledge. However, since incidental learning relies on triggers, the challenge is that the learning process is fragmented and incoherent (Kukulska-Hulme et al., 2015; Paletta et al., 2013). Observational learning occurs when people learn through observation and model activities and outcomes from their environment. As a consequence, they can learn behaviour before actually performing it (Krath et al., 2021). Additionally, this thesis builds upon the

work of the existing serious games of *Change Perspective* and the *Dilemma Game*. Thus, continuing in the format of a serious game is a logical step.

To our knowledge, no serious game about the combination of diversity and (international) student inclusion exists yet. Barrera Yañez et al. (2020) reviewed serious games representing gender equality, but gender is only one aspect of diversity. Serious games representing multiple aspects of diversity seem to be missing.

To focus on the second part of the combination, serious games targeted at (international) student inclusion are lacking, let alone those that focus on the Netherlands in particular. Rienties and Tempelaar (2013) and Rienties et al. (2014) researched international student inclusion in the Netherlands and their results indicated that "some groups of international students experience considerable personal-emotional and social adjustment issues" (Rienties & Tempelaar, 2013, p.188). This finding was applicable to Eastern European and non-European students, and in particular for students from Southern and Confucian Asia. Concerning facilities to support student inclusion, universities often offer a Buddy Program. Nevertheless, these programs often have limited spots. Furthermore, as a participant, one is dependent on their Buddy and the type of activities offered. These constraints are not present in a serious game (Nilsson, 2019).

Although there exist some games about (international) inclusion (see for example Paletta et al., 2013), this is still a growing field. In short, there is no serious game about the inclusion of different groups at a Dutch university yet. The *Dilemma Game* mentioned before is a step in the right direction. However, it is a card game, which limits its flexibility in teaching, since the game cannot adapt to the user in real-time, while a digital serious game could.

1.2 Contribution

As mentioned before, intergroup biases are problems, especially when a 'we-vs-them' atmosphere dominates. Higher education institutions still fight against inequity and ignorance instigated by prejudices and biases. This may be because such institutions traditionally have been very exclusive, instead of inclusive (Dwyer & Smith, 2020). To counteract this opposing feeling and stimulate diversity instead, a solution could be to encourage intermingling between these different groups. By developing *Diverse Perspectives*, this research attempts to create a better understanding of how such a solution could be shaped. It intends to do so by investigating the combination of diversity and (international) student inclusion.

Moreover, it aims to contribute to the field of inclusive serious games targeted at (international) students by providing a serious game that could inspire future research. The underlying message of *Diverse Perspectives* is that it should teach users to see things from another point of view, especially the underrepresented ones. Thus, our vision is to raise awareness of diversity topics and familiarise the player with the unfamiliar.

1.3 Outline

The rest of the thesis is structured as follows. In Chapter 2, the Related work is presented. Chapter 3 describes a high-level overview for the Design of the game. The Implementation is defined in Chapter 4, whereas the Method is explained in Chapter 5. Chapter 6 reports

the Results and Chapter 7 depicts the Discussion. Lastly, Chapter 8 presents the Conclusion. Supplementary material can be found in the Appendices.

Chapter 2

Related work

This chapter discusses what serious games are, the theories behind it and touches upon the concept of storytelling and visual novels. Moreover, it describes similar serious games in the field and elaborates on the topic of diversity.

The reviewed literature was retrieved through Google Scholar in incognito mode. The following keywords were entered:

- Keywords ([...] replaces the word(s) in italics)
  - *Serious games*: [...] for international students, diverse/inclusive [...], racial bias [...], narrative [...], [...] cultural integration, [...] student integration, buddy programme, [...] learning, game-based learning, gamification, [...] branching, [...] storytelling, PENS/GEQ questionnaire
  - *Diversity/Inclusion*: intersectionality, scale to measure [...], promoting [...], [...] workshops, [...] seminars, [...] initiatives, universality diversity scale, character representation in video games, inclusive character design
  - *Social psychology*: need to belong, group dynamics, similar attraction theory, contact theory, social identity theory, information decision-making theory, intergroup bias, self-determination theory, flow theory, cognitive load theory

Snowballing caused searches for specific papers. Snowballing sources consisted of literature from several courses, including *Serious Gaming (INFOMSEGA)*, *Advanced Cognitive and Social Psychology for HCI (INFOMCSP)* and *Advanced Introduction to Gender Studies (MCRMV16016)*.

If no specific paper was targeted, the strategy for choosing papers was checking if the article was relatively recent. That is, it was published after 2016. From that pool, the title, abstract and number of citations were checked for relevance. If the abstract seemed promising enough, the paper was read and included in the literature review where applicable.

2.1 What are serious games?

This section explains what serious games are, how they differ from entertainment games, gamification and game-based learning, which different domains exist, and the theories behind serious games.

2.1.1 Difference between entertainment games, serious games, gamification and game-based learning

A serious game is not the same as an entertainment game. Examples of entertainment games are *Life is Strange*, *Animal Crossing* and *Mario Party*. A serious game is a digital game that has at least one additional goal besides entertainment. These additional goals are called characterising goals. Examples of such goals are learning mathematics or choosing a healthier lifestyle. In other words, entertainment is not the primary goal of these games. People tend to be less interested due to the label 'serious', yet serious and fun do not need to be mutually exclusive. Rather, the degree of seriousness and fun could be seen as a trade-off. Serious games have the challenge to find the right balance between entertainment and the educational message (Laamarti et al., 2014).

Early examples of serious games are *The Oregon Trail* and *Where in the World is Carmen Sandiego?* (Naul & Liu, 2020). Created in 1974, *The Oregon Trail* puts the player in the role of a wagon leader in 1848. The objective of the game is to survive and to guide the group from Missouri to Oregon. By following this trail as a pioneer, the game teaches its players the history around it (Bouchard, R. Philip, 2017). Made in 1985, the goal of *Where in the World is Carmen Sandiego?* is to catch the art thief Carmen Sandiego. Players can do so by travelling to the correct destination, with the help of clues. These clues contain general knowledge about the world, thereby the game intends to let the players obtain this new knowledge (Chabebe Rivera et al., 2022). A more recent example of a serious game is *Peacemaker*, made in 2007. Here, the player learns the complex situation between Israel and Palestina through the eyes of the Israeli Prime Minister or the Palestinian President (Belman & Flanagan, 2010).

Related terms to serious games are 'gamification' and 'game-based learning'. In gamification, game elements (e.g. rewards, a leatherboard) are added to existing non-game learning activities. A well-known example is the language learning platform *Duolingo*. Although language learning is not a game in itself, the process of doing so is made more engaging by adding points, lives and trophies (Shortt et al., 2021). In game-based learning, the learning activities are shaped in the forms of game content and play. *Kahoot* is an example of game-based learning, where players can test their knowledge by competing in customisable quizzes that capture the learning materials. Thus, serious games are products that employ game-based learning techniques (Wang & Tahir, 2020). In other words, game-based learning and serious games rely both on complete games, while this is not the case in gamification (Krath et al., 2021).

2.1.2 Different domains of serious games

Serious games intersect with several other disciplines, such as psychology, pedagogy and Artificial Intelligence (AI). Therefore, they can be applied to several domains, such as education, training simulations and health care. Laamarti et al. (2014) created a taxonomy to

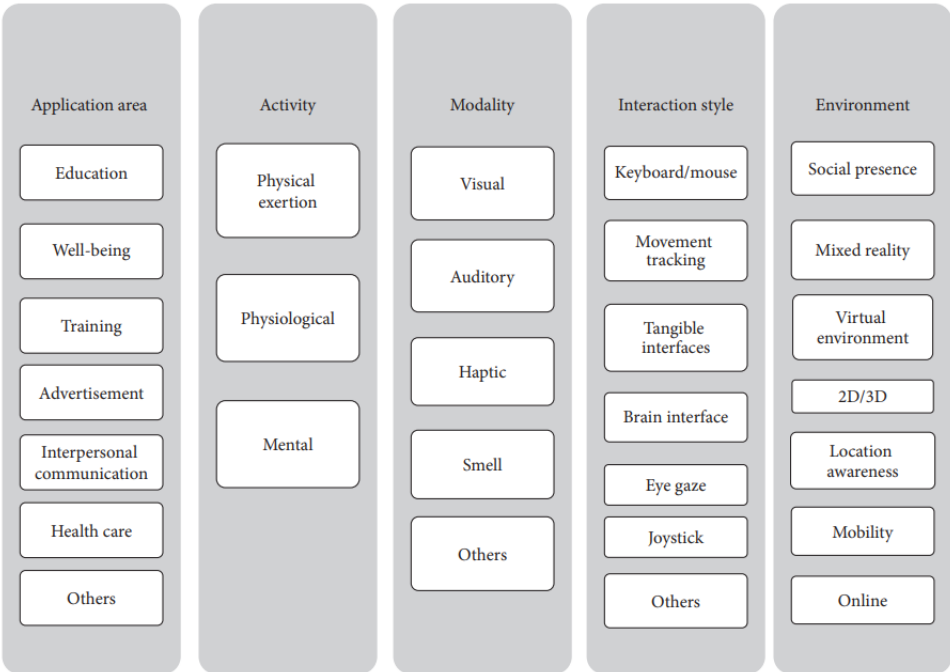


Figure 2.1: An overview of the classification of serious games distributed along five categories, created by Laamarti et al. (2014). A game could consist of multiple sub-categories per category.

.

classify serious games, based on their characteristics. This can be seen in Figure 2.1. Their taxonomy consists of five categories: 1) Application area, 2) Activity, 3) Modality, 4) Interaction style, and 5) Environment. Note that a game could consist of multiple sub-categories per category. First, the application area points out the relevant domains a serious game could apply to. Next, the activity denotes the type of input that is needed from the player to play the game. Third, the modality is the communication channel from the game to the player. The interaction style refers to how this player input occurs and what devices are involved. Lastly, the environment defines the platform of the game. That is, how the surroundings of the game are supported.

To classify *Diverse Perspectives* in the taxonomy, the application area could be 'social game', as the game intends to teach players about the social dynamics of student groups at UU. *Diverse Perspectives* is not an educational or a training game, as it does not focus on educational subjects such as mathematics or language learning, or simulation purposes such as surgeries or emergency situations. The game neither belongs to the well-being area or health care area, since it is not a game about (mental) health. Nevertheless, these are related areas, due to the negative (mental) health consequences of exclusion and intergroup biases. Lastly, *Diverse Perspectives* is not an advertisement game, as it has no advertisement goals. However, the 'social game' subcategory is not present in Figure 2.1. Therefore, it could fall under the subcategory of 'Interpersonal communication', since communication between the player and other characters plays a central role, or under the subcategory of 'Others'.

As *Diverse Perspectives* is a digital game and therefore playable on a computer or a laptop,

there is no physical or physiological effort needed, besides cursor movement. Thus, the type of activity that is involved belongs to the 'Mental' subcategory, since *Diverse Perspectives* only requires cognitive capacities. Additionally, the modality of the game is 'Visual', as the game can be played via a computer or laptop. These devices contain the 'Keyboard/mouse' interaction style, as no other equipment is necessary to play the game. Because *Diverse Perspectives* has the format of a visual novel, there is no social presence, since it is played singleplayer. Furthermore, it does not belong to the 3D, mixed or virtual reality environment, as the game is represented in 2D. Moreover, it does not need the player's location and the game is neither playable (yet) on a mobile phone nor online. Hence, the environment is only '2D'.

2.1.3 The theory behind serious games

Krath et al. (2021) researched the effects and outcomes of serious games. They found 3 types of outcomes: 1) motivational and affective, 2) behavioural, and 3) (cognitive) learning. Motivational outcomes represent the fact that playing games is perceived as motivating and it explains the reason why people engage in certain behaviours. Valence, which is "an evaluation of the subjectively experienced state" (p.2), is seen to elicit motivation. Secondly, positive behavioural outcomes, such as engagement, teamwork and performance improvements are found with serious games. Lastly, various cognitive learning outcomes are associated with serious games. Examples are improvements in critical and creative thinking, obtaining knowledge and understanding of materials and perceptual skills. Nevertheless, it is suggested that these outcomes in the cognitive domain need to happen in combination with motivational and learning outcomes to be beneficial. Furthermore, the authors performed a systematic literature review to research the theories used in serious games. Although they remark that no individual theory is able to explain how serious games work, we will highlight three well-known theories: Self-Determination Theory (SDT), flow theory and cognitive load theory.

2.1.3.1 Self-determination theory

Motivation exists on a continuum. At the one end, people are completely intrinsically motivated. At the other end, people are completely externally motivated. Behaviour is intrinsically motivated when it is driven by internal rewards. In other words, just performing the behaviour is rewarding in itself. In contrast, behaviour is externally motivated when performing it serves a higher goal or an ulterior motive, for example obtaining external rewards or avoiding punishments. While intrinsic motivation is understood to be universal and innate, it can be developed in a social context. Yet, the boundary between intrinsic and external motivation is not always clear-cut, as it could be difficult to determine what motives can be seen as intrinsic or external (Kruglanski et al., 2018).

Thus, it is most desirable that people are intrinsically motivated to learn. In order to achieve intrinsic motivation, SDT states that three psychological needs should be fulfilled when people interact with their environment. The first need is *autonomy*. This need is met when people see themselves as the origin of their own behaviour and that they can choose their actions in such a way that they align with their own values and interests. The second need is *competence*. Competence is satisfied when people experience the feeling that they are skilled enough to effectively perform their activities. *Relatedness* is the third need and is accomplished when people feel a sense of connectedness and belongingness with others (Deci

& Ryan, 2008). Intrinsic motivation in serious games can be increased in several ways, such as avoiding boredom and stress by adapting to the difficulty level to the user, using persuasive strategies and using words of affirmation (Wouters & Oostendorp, 2017).

2.1.3.2 Flow theory

Flow theory presents the concept of flow. When people are in the flow, they are in such a state of concentration that they are fully focused and absorbed in the activity. Typically, they feel at the peak of their abilities, effortlessly in control and forget about their current problems and the passing of time. Flow makes experiences genuinely pleasant and this state is reached when the activity is neither too simple nor too easy. In other words, the difficulty level of the activity and the skill level of the person are in perfect balance. If a task is too difficult, people risk experiencing anxiety, failure or maladaptation. If a task is too simple, this can invoke boredom, lack of improvement and deterioration. Dörner et al. (2016) visualised this in a graph, which can be seen in Figure 2.2. Note that it is possible to have different degrees of flow, varying from low, medium and high. Eight major components of flow are: 1) clear goals and immediate feedback, 2) balance between task difficulty and skill, 3) merge of action and awareness, 4) focused concentration, 5) sense of control, 6) loss of self-consciousness, 7) time distortion and 8) self-rewarding experience (Csikszentmihalyi, 1990).

As serious games have the twofold goal of being both effective and engaging to play, the concept of dual flow is characteristic and unique for these games. That is, players should reach their flow state both in terms of effectiveness and engagement. Furthermore, each player has their own flow channel, which means that the game should be designed in such a way that it allows every player to enjoy their personal flow. Disruptions of this flow path should be avoided. Since serious games assume that the player's skill will improve over time, the task difficulty needs to be adapted accordingly to let the player stay in their flow channel. It could happen that the task difficulty is set too high and in that case, the player might return to their flow when they improve. Otherwise, the difficulty level should be adapted again, to avoid players moving to the anxiety zone (Streicher & Smeddinck, 2016).

2.1.3.3 Cognitive load theory

Cognitive load theory proposes that the human brain has a limited amount of working memory. This limit indicates that most people can remember between three and seven separate pieces of information at the same time. Information overload occurs when this limit is exceeded. This should be avoided as it reduces the effectiveness of the learning process. There are three types of cognitive load: 1) Intrinsic Cognitive Load (ICL), 2) Extraneous Cognitive Load (ECL), and 3) Germane Cognitive Load (GCL). ICL consists of all information that is relevant to the learning process, such as educational content. The load is determined by the complexity of the information. ECL is created by learning-irrelevant processes, such as lay-outs. In other words, the load is increased by distractions. Lastly, GCL represents the linking of new information to knowledge that is already present in the long-term memory. GCL is desirable, as it denotes the retention of new information.

As mentioned before, people have a limited working memory capacity. When there is too much ICL and ECL, there is not enough working memory left to handle the GCL. That is, the movement of new information from the working memory into the long-term memory does not occur. Because people have little control over the complexity of information, there

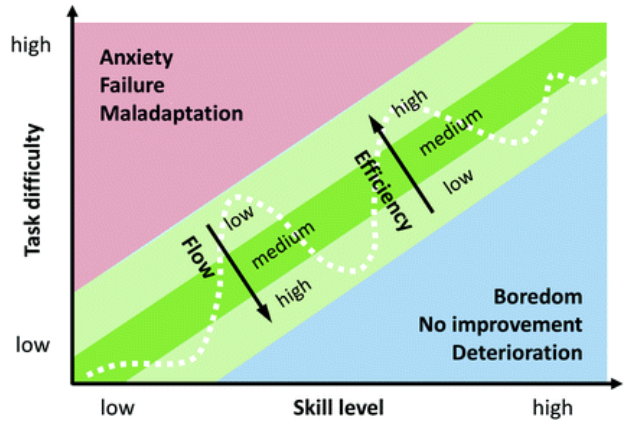


Figure 2.2: The dual flow in serious games, created by Dörner et al. (2016). The original figure contained a typo ('maldadaption'), which has been corrected here. When the task difficulty is too high, players move in the anxiety zone. When the task difficulty is too low players move in the boredom zone. When the task difficulty and the skill level are in balance, the player is in the flow. This flow can vary in degrees of low, medium and high.

.

is not much room to reduce the ICL. However, people do have control over the ECL and thus it should be reduced as much as possible (Loveless, Becton, 2022; Sweller, 2010, 2011). Serious games could play a role in reducing ECL, and therefore make educational content easier to learn. However, other studies suggest that serious games contribute to the ECL, as the educational messages are wrapped in the game. Hence, it is still a debate if serious games decrease the ECL or rather increase it (Krath et al., 2021).

2.1.4 Pedagogical learning theories

Because serious games have an educational element, pedagogical learning theories play an important role in the game design. This subsection discusses two models that were applied in the design of *Diverse Perspectives*: the Attention, Relevance, Confidence and Satisfaction (ARCS) model and the Eight Learning Events Model (8LEM).

2.1.4.1 ARCS

In the pedagogical process, motivation is deemed essential to encourage and maintain learning. In order to stimulate this motivation, ARCS can be used. This is a motivational design model. Its acronym contains four requirements: 1) Attention, 2) Relevance, 3) Confidence, and 4) Satisfaction. These should all be met for people to become and stay motivated.

The first condition is attention, which is a prerequisite for learning. A strategy to obtain the learner's attention is to let them interact with the materials. Although getting attention is the first step, the challenge lies in maintaining it. The teaching materials should find a balance between understimulation and overstimulation, to prevent both boredom and anxiety. In *Diverse Perspectives*, the format of the game is used to attract the learner's attention, as playing a game might be more appealing due to its participatory nature than passively reading literature or following a lecture.

The second condition is relevance. If the teaching material is perceived as relevant, this will be beneficial for the learner's motivation. An approach to increase the relevance is to explicitly state how the material can be used in the learner's future. In the case of *Diverse Perspectives*, the story represents student life in Utrecht. At the end of the game, the player receives a recap with take-away messages that could be used as guidelines in the future. Moreover, the game intends to fit in the learner's world. This is because the story revolves around being a student in Utrecht, which aims to relate to the target group.

Confidence, or expectancy for success, is the third condition. That is, the game should create the impression that a certain degree of success can be achieved if the learner puts some effort in it. A method to stimulate confidence is to allow the learner to learn new knowledge in safe environments, but let them apply this knowledge in realistic situations. *Diverse Perspectives* provides this safe, but realistic environment. Because it is a game, the choices the player makes only affect the game world. However, their choices do matter and have consequences in that particular world. Since the game is choice-based, they receive plenty of opportunities to shape the outcome, if they put enough effort in thinking and reflecting.

The last condition is satisfaction. This means that the learner should have a positive feeling about their achievements. Satisfaction can be met by giving appropriate rewards. A strategy to increase satisfaction is to give the learner the opportunity to apply their knowledge in real life. In *Diverse Perspectives*, the reward is reaching the good ending out of the three endings, while the punishment is getting the bad ending. Thus, the game implements a reinforcement strategy (Keller, 1987; Li & Keller, 2018).

2.1.4.2 8LEM

Another pedagogical model is the 8LEM. The model connects the student's demand and the teacher's supply. That is, the activities are complementary and interdependent. The eight learning events are: 1) observation/modelling: learning through observation and imitation, 2) reception/transmission: learning by receiving information or advice, 3) exploration/documentation: learning via free investigation, 4) self-reflection/co-reflection: learning via reflecting on one's skills and knowledge, as well as the learning process, 5) debate/animation: learning through social interactions, 6) creation/encouragement: learning by creating new content or applying familiar content, 7) experimentation/reactivity: learning through environment manipulation and observation of effects, and 8) exercising/guidance: automating skills via guidance and corrective feedback.

A core principle of 8LEM is that the variety in methods is beneficial for the learning process, such as the development of skills and maintenance of motivation. In other words, using different teaching approaches is preferred above continuously using the same method (Verpoorten et al., 2007).

Diverse Perspectives uses several methods from the 8LEM. For example, it applies method 2), as the narrative of the story conveys information that can be useful to the player. Likewise, it utilises method 3), since the player is free to make their own choices to discover what happens. Furthermore, the game implements method 4), because it requires the player to reflect on the learning content before making a choice. Lastly, the game employs method 5), as the player interacts with other Non-Player Character (NPC)s and therefore receives various viewpoints on the same topics.

2.2 Storytelling

Diverse Perspectives relies on a story to convey its educational message. Stories go well with learning, as they are easy to follow and to comprehend, since they have a familiar format, while still leaving room for interpretation (Naul & Liu, 2020). A story consists of a narrative and a progression. Narratives can be told via a string of pearls, where each pearl denotes an episode and they are connected via strings that can contain cut scenes. Alternatively, it can be created by the players themselves in a story machine, for instance in *Minecraft*.

The narrative is told by using common techniques that are likewise used in novels or movies. However, unlike the traditional storytelling media, the narrative of a game does not need to proceed in a linear way, but can rather progress through different branches and alternative endings - which is the case of *Diverse Perspectives* - or by using foreshadowing and backstories. Therefore, the game can have a dynamic and interactive story. Nevertheless, this creates the *Narrative Paradox*, meaning that there is a conflict between narration and interactivity, as every form of interaction breaks the narrative curve. To keep the player engaged and in the flow, their interest curve should gradually increase over time until the end of the story is reached, with local peaks in between. These local peaks can come after the end of each episode, for example, or in the form of cut scenes (Göbel & Wendel, 2016). Moreover, the dialogue between characters is of vital importance, especially in visual novels. Even though the dialogue differs from the dialogue in a book, it should still be natural, interesting and sensible, as it can make or break the game (Mildner & Mueller, 2016; Schell, 2015).

2.2.1 Different types of narratives and progression

The story of a game can unfold in various ways. Carstensdottir et al. (2019) describe six different types of 'structures' (narratives) and four different types of 'progression mechanics'. Structures are graph-based layouts of the narrative, while progression mechanics denote the rules that allow players to move along this graph. In other words, the manner how events are presented to the player depends on the structure. The six different structure types are: 1) linear, 2) branching, 3) foldback, 4) broom, 5) hidden, and 6) opportunistic story.

Linear stories have a single start and a single ending. There is only one path to be followed, thus all events are presented in a fixed order. An advantage of these linear stories is that the storytelling can be controlled and is cohesive, while a disadvantage is that they are not suited for replayability.

Branching stories have multiple endings, as multiple paths can be followed to reach them. Each path differs from another, hence there are multiple narratives to be told. This serves well for replayability, however a drawback is that branching stories are complex to build, as the content of the story should stay consistent and the number of branches could grow exponentially depending on the number of choices (Schell, 2015). Therefore, reducing the number of choices or the number of outcomes to a manageable amount is important to maintain the story feasible. *Diverse Perspectives* used this type of structure as well. This is because branching allows players to reflect on the benefits and limitations of certain choices and to form connections between them (Øygardslia et al., 2020).

Foldback stories have isolated branches. Players can make choices to choose a certain path,

but in the end this path converges back to a single event. A benefit of this is that the narrative can contain some variation, while still controlling for the number of branches. Nevertheless, players could notice this limited variation in replays.

Broom stories follow mainly a linear structure, while having multiple branches leading to multiple endings. They differ from branching stories, as the majority of the story is presented in a fixed order. Only at the end, multiple paths become possible. This structure type has the same strengths and weaknesses as the foldback structure.

Hidden stories contain pieces of narrative that are broken up and distributed among the game. These stories are often unlocked by encountering hidden objects. The player is supposed to craft the story together from these separate chunks, the narrative components lack explicit paths between them. An advantage is that this type of structure offers a high level of control to the player in terms of how they want to experience the narrative. A disadvantage is that there is often a finite amount of content, which limits replayability.

Opportunistic stories contain the most dynamic structure. Depending on the player's actions, countless variations are possible of how the events are presented. It can be understood as an 'open' narrative. That is, infinite numbers of events can take place in any order. While many replays are possible in this structure type, its complexity is significant.

Moreover, the four different progression mechanics are: 1) progress through choice, 2) progress through scripted scenarios, 3) progress through discovery, and 4) progress through in-game systems. Each mechanic not only conveys the intent of the game designer, but constraints the player's behaviour as well.

Progress through choices allow players to continue the story by making direct choices. Often, the game focuses on the consequences of each choice, thereby encouraging the player to reason about the path that will follow once they made a certain decision. A choice can be made in two ways: selection or performance. In the first case, the player chooses an option explicitly, such as via a menu or dialogue box. The player is unable to continue the story until they make a decision. In the second case, the player 'performs' the choice, which is more implicit and often relies on environmental elements. For example, walking through a certain door. Progression through choice selection is used in *Diverse Perspectives*, as the player is able to make decisions through dialogue boxes. Additionally, it is very common to feature branching stories with progression through choice.

Progress through scripted scenarios occurs to scripted events where players have to perform a specific pre-scripted action to continue the story. Often, these events are displayed in interactive cut-scenes where players have a limited time to make a choice by pressing a certain button. Which button they pressed and even their reaction time could decide which outcome they receive. Failing to do so could either impact or prevent progression of the game.

Progression through discovery means that players have to find story content to proceed the story or could even unlock additional parts. That is, players should search, explore and investigate to make progression in the game. A progress bar could be provided to indicate the player of their progression status.

Progression through in-game systems requires players to interact with in-game systems to continue the story. An example is task completion, where players have to complete tasks in order to progress. Another example is resource management. That is, players have to reach certain skill levels to proceed. Lastly, progression through combat is another instance, where the outcome of battles could influence the progression of the story.

2.2.2 Game design principles for visual novels

According to Øygardslia et al. (2020), visual novels have three elements: 1) narratives containing mainly text, dialogue boxes, backgrounds and character sprites, 2) showing illustrations to the player at central points in the narrative, and 3) branching structure with multiple endings, depending on the player's choices. Furthermore, they formulated six game design principles for visual novels. These design principles are desired to be taken into account in *Diverse Perspectives*.

The first design principle states that educational content should be easily accessible. Moreover, the choices should be designed in such a way that the player wants to explore for more information that is related to the learning goals. The game intended to do this by presenting the learning material in the form of a choice, where the answer influences the outcome of the story.

The second design principle defines that narratives should be used to encourage reflection by highlighting certain topics in different contexts or from different viewpoints. *Diverse Perspectives* aimed to implement this by presenting the player different scenarios covering various topics. Different viewpoints are conveyed through the perspective switch.

Design principle three explains that characters should be developed carefully to promote player identification. Additionally, designers should collaborate with target users to ensure that marginalised groups are represented accurately. *Diverse Perspectives* intended to do so by including a diverse cast of characters. More on character design can be found in section 4.2.4.

The fourth design principle describes that players should be stimulated to pursue the paths they find the most interesting. Likewise, players should have the opportunity to learn from the consequences of their actions, both the good and the bad ones. Since in *Diverse Perspectives* each choice results in a different consequence, this invites the player to pursue and forge their own path tailored to their interest and preferences. The ending they receive serves as feedback: if players perform good actions, they are likely to obtain the good ending and vice versa.

Design principle five states that the narrative should be structured, yet creative. The learning goals should be clear, but their implementation should spark the player's reflection, reasoning and decision-making. *Diverse Perspectives* calls upon the player's cognitive skills, as their decision-making shapes the narrative. While each player receives the same core narrative, its structure contains several branches that result in different stories.

The sixth design principle describes that the game should have clear goals and assessment strategies. This could include 'stealth assessment'. That is, the assessment is integrated in

the gameplay. Broadly speaking, there are two categories: 1) letting the player apply their knowledge in order to proceed, and 2) providing feedback to the player if they choose the correct answer. *Diverse Perspectives* used this stealth assessment, as it tracks the player's choices and interactions with the NPCs via a hidden score system. Each choice regarding a diversity dilemma could change the score and the total score in the end determines which outcome the player receives. This score system is further elaborated in section 4.2.3.1.

2.3 Serious games in the domain of inclusion and diversity

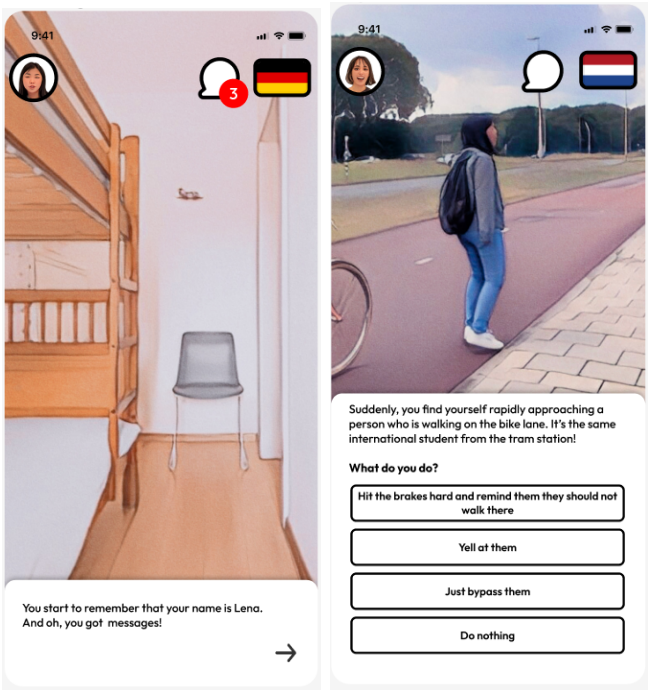
Various related serious games are discussed in this section, such as *Change Perspective*, *MASELTOV* and the *Dilemma Game*.

2.3.1 Change Perspective

Change Perspective is a serious game made in the *Serious Gaming (INFOMSEGA)* course in 2022 at UU. The game is targeted at both international and Dutch students, where it aims to promote intermingling between these two groups through cultural integration. It was noticed that these groups do not interact very frequently, while mainly international students would like to have more interaction with local students to feel at home in their host country. *Change Perspective* is a narrative-driven game with multiple endings, which means that the player's choices affect the outcome of the story. Although the game is singleplayer, players play from both the First-Person Perspective (1PP) of a Dutch student (Kim) and a German student (Lena). During the story, these perspectives switch multiple times, to let the player experience the same situation from different viewpoints. This is because the game intends to emphasise that the responsibility of inclusion lies with the Dutch students as well, and not only with the international students. For international students, *Change Perspective* allows them to learn about a new culture of the host country in a 'safe' space and to avoid offensive or embarrassing behaviour in real-life situations. For Dutch students, the game allows them to develop a more open-minded attitude by teaching them the value of interacting with international students and creating empathy for their position.

A prototype of the game is made in Figma and can be played here[1]. *Change Perspective* is a 2D game that can be played on any device that has access to Figma. No other equipment is needed, as it relies on a point-and-click system. The layout of the wireframes are based on a mobile phone, as there were ideas to potentially make the game into a mobile application. Figure 2.3 displays screenshots of the game through the 1PP of the German student Lena (a) and the Dutch student Kim (b). On the top left, the player's avatar (the character they are playing) is shown in a small circle. On the top right, the flag of the country of that particular avatar is displayed. This is to emphasise through whose eyes the player is playing. The speech bubble, representing the avatar's phone, is located to the left of the flag. A number indicates that the avatar has unread messages. The player can open these by clicking on the speech bubble. These messages are part of the story as well and serve as an extra channel of telling the story besides the dialogue boxes. Figure 2.3b represents a situation where a choice between four options should be made.

[1] Change Perspective prototype



(a) (b)

Figure 2.3: Screenshots of *Change Perspective* from the perspective of the German student Lena (a) and the Dutch student Kim (b). The player interacts with the game via a point-and-click system (Backx et al., 2022).

The story of the game focuses on a Dutch *hospiteeravond* (i.e. *hospi*). This is an evening where you and other people are invited to visit a potential room. The goal of the event is to see who is the best tenant candidate for the room. Usually, the residents of the house decide who matches best, but it is also possible that the landlord has the final say. Both Kim and Lena are invited to a *hospi*, as they are desperately looking for a room in Utrecht. Depending on the player's choices in the game, they can achieve one of the three endings. In the good ending, the *hospi* goes well, since Lena and Kim befriended each other and they are both chosen to be the new roommates in the house. In the bad ending, the *hospi* does not go so well, as Kim and Lena do not go along and therefore do not make a good impression. Lastly, in the neutral ending Kim and Lena are both liked by their potential roommates, but in the end only Lena receives a call that she gets the room.

The game has been evaluated through multiple focus groups and improvements consisted of a clearer switch between perspectives, a short recap of important material and the development of more episodes, so that the game could cover more situations. Yet, Change Perspective was generally perceived as positive and is deemed to have the potential to improve cultural integration between students. Therefore, it serves as a backbone for this thesis (Backx et al., 2022).

2.3.2 MASELTOV

2.3.2.1 Development of MASELTOV

Mobile Assistance for Social Inclusion and Empowerment of Immigrants with Persuasive Learning Technologies and Social Network Services (MASELTOV) was an European project that was active from January 2012 until March 2015. The serious game was targeted at people that migrate to or within Europe to support them during their integration process and mitigate the risk of social exclusion. Playable as the MASELTOV App (MAPP) on Android mobile devices, it was a mixed reality game that offered learning services in an engaging way. Examples of such services were the development of communication skills, incidental learning of the host country's language and culture and supporting employability. As mentioned in 1.1.2, a challenge for incidental learning is that it often occurs in short episodes and the process is therefore fragmented. To overcome this challenge, feedback and progress indicators should be used to help learners stay motivated and take responsibility for their own learning process.

The idea of the game was that it allowed the 'scaffolding' of learning. That is, the more a player plays the game, the less support is required from the game to handle situations in daily life, until the support is no longer required and can fade into the background. Additionally, MASELTOV uses a 'freemium' model, meaning that players can access premium content based on their time and effort put into it, instead of accessing it via in-game purchases. Reasons for developing MASELTOV instead of traditional educational materials were that it would be more appealing to some players due to its informal and playful nature, thereby reaching a broader audience. Moreover, the format of a game could highlight the cultural differences between the player and their host country, in the form of an in-game narrative. Likewise, to stimulate understanding between different cultures, MASELTOV included a 'dimension flip'. This allowed the player to perceive situations from the perspective of two separate opposing cultures and how these differences might conflict with each other. The concept of this dimension flip is similarly used in *Change Perspective*.

A feature that was included in MASELTOV were the language learning exercises that focused on daily tasks. These exercises conformed with the Common European Framework of Reference for Languages (CEFR) at A1 and A2 level, and were peer reviewed. Another feature was the mobile navigation tool. This tool was similar to other navigation tools such as Google Maps, and it could provide the player directions, points of interests and services (e.g. libraries, hospitals). Furthermore, the tool supported routes for pedestrians and public transport. Once the player selected a location or was in close proximity to it, a language learner exercise could be triggered. A third feature was the profile system. Information of the player's personal information and their learning process was stored in the game to provide personalised learning and recommendations. The fourth feature was a geo-social radar, which allowed the player to find nearby volunteers who could help them in case they had a problem. Lastly, the fifth feature was the TextLens. This feature enabled learners to convert a photo of a sign to text and have this translated to their native language. Moreover, images and text could be uploaded to a forum for questions and discussions. While *Diverse Perspectives* will not contain the aforementioned features, MASELTOV still serves as a relevant example.

When developing MASELTOV, a user-centred design approach was followed. That is, the involvement of target users in the early design stages. Several semi-structured interviews and focus groups were held to gather user feedback and construct scenarios. As contacting

all European immigrants was unfeasible, three large target groups were identified to take the cultural diversity into account. These were: 1) North African immigrants that speak Arabic, 2) Turkish-speaking Turks, and 3) Latin Americans that speak Spanish. These three groups differ from Western cultures and are therefore considered suitable as target users (Paletta et al., 2013).

2.3.2.2 Evaluation of MASELTOV

As said in the previous section, the MAPP contained language learning tasks, image-to-text translations, location-based and personalised recommendations, game-based cultural learning and a social network. The language lessons consisted of 6 'modules', each focusing on their own theme, for example travel or healthcare. Each module contained 3 to 4 'lessons', a 'situations' page, which had audio and written text snippets, and an 'emergency vocabulary'.

MAPP was evaluated in three countries by the three target groups. Each group had a different evaluation focus. The division was as follows: 1) Latin Americans in the United Kingdom. They focused on language learning and the game, 2) North Africans in Spain. This group concentrated on communication, and 3) Turks in Austria. They fixated on the navigation part. In general, the app was evaluated positively. The translation and language learning tools were highly popular and participants mentioned that these services were beneficial in their language learning process. Furthermore, the navigation services were also highly appreciated, suggesting that travel and communication play a vital role.

The social integration part of the game was received with mixed results. Players liked that these services allowed them to interact with others, but were concerned how they could protect their privacy. Likewise, the serious game had mixed results as well. Some players enjoyed it, while others did not have any interest in playing a game or found it confusing. It seemed that having immediate solutions to everyday problems in the form of tools and services was valued more highly than learning through games. Lastly, players remarked that a high degree of mobile literacy was needed to use MAPP and that many features of the app needed a constant internet connection. Another suggestion consisted of giving players longer-term learning journeys, instead of fragmented episodes, so learners would be able to reflect on their learning process (Kukulska-Hulme et al., 2015).

2.3.3 The Dilemma Game

The *Dilemma Game* was developed in 2019 by the members of the Nanophotonics Section from the Debye Institute for Nanomaterials Science of UU. It is a card game that contains 40 dilemmas. A link to all the cards of the renewed version of the game can be found here[2]. Although it is stated that the game is "especially relevant for employees in the (higher) education sector", students should play it as well, for example on introduction days. The *Dilemma Game* durates at least 60-90 minutes and should be played in groups of 2-7 people. There are no right or wrong answers, instead the goal of the game is to stimulate discussions between the four options offered at each dilemma. In other words, each dilemma serves as a conversation starter on different issues and players are encouraged to find solutions together. Judging or criticising is not desirable, rather players should ask for clarification to understand the choices of other players.

[2]https://www.uu.nl/en/dilemma-game

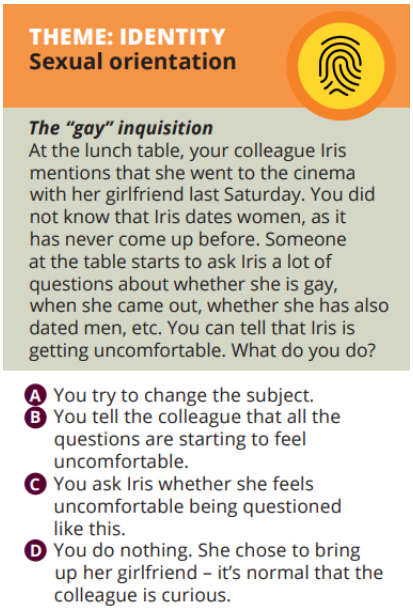


Figure 2.4: A dilemma from the *Dilemma Game* (Equality, Diversity & Inclusion (EDI) Office of Utrecht University, 2022). This dilemma falls under the 'Identity' theme and the 'Sexual orientation' subtheme. Each dilemma has four options and players should share with each other which option they favour and why.

The dilemmas are divided into different themes, all related to diversity and inclusion. There are five themes, where the first two are divided into subthemes: 1) identity, with subthemes 'women and men', 'transgender and non-binary people', 'sexual orientation' and 'sexually abusive behaviour', 2) cultural differences, that has the subthemes 'religion', 'ethnicity', 'culture' and 'language', 3) privacy, 4) hiring, and 5) disabilities. An example of a dilemma is shown in Figure 2.4.

However, printing a real card game for all students at the introduction is unfeasible and thus a digital version is preferred to reach its full potential (Equality, Diversity & Inclusion (EDI) Office of Utrecht University, 2022). Hence, *Diverse Perspectives* aims to continue in a similar direction by taking this game as a basis for the content and the design of the choices that will be presented in the story.

2.3.4 Other related serious games

Fonseca et al. (2021) created *Secrets of the South*, which is a serious game targeted at children and teenagers between 10-16 years old that live in Rotterdam. The game makes use of the player's location and challenges them to visit people or places where they would normally not go to. Therefore, the goal of *Secrets of the South* is to enhance the player's engagement and integration with their own neighbourhood through meaningful social interactions the game creates. Meaningful in this context means "an overall enjoyable experience for a player, and enjoyable or neutral for individuals involved in the gameplay" (p.1).

Moreover, the authors present design recommendations for meaningful social interaction. Results indicated that *Secrets of the South* made social interactions happen and that these interactions had a mainly positive effect. Likewise, players worked together in all kinds of

challenges and that the level of collaboration was intense for the majority of the challenges. In other words, the game allowed players to socially interact with others and their neighbourhood. Although the integration aspect of this game is more locally focused, it could still serve as relevant inspiration for *Diverse Perspectives*.

Passage Home is a serious game that addresses racial bias. A challenge is to correctly understand the attitudes towards racial and ethnic identities, as an accurate representation of it is vital for the serious game to be an effective learning tool against bias. *Passage Home* aims to investigate how one's racial bias influences their interpretation of the story in the narrative game. The story involves Tiffany, a Black student, being falsely accused of plagiarism by Mrs. Smith, her white teacher. The player's avatar is Tiffany, represented in 1PP, while Mrs. Smith is a NPC. The player is told that Tiffany is a hard working student and put a lot of effort into creating an original essay. Yet, Mrs. Smith believes that Black students do not have the capacity to write such an essay and thus concludes that Tiffany must have committed plagiarism. During the game, the player receives prompts to select their thoughts, feelings and actions they wish to perform.

Passage Home is implemented in two versions: a 3D representation and a text-based representation. Results indicated that the player's ethnic identity and their attitude towards race influenced their interpretation of the narrative and their game experience. The findings suggested that players who are not aware of racial bias in the real world, may fail to do so as well in the game. Additionally, they may fail to recognise the importance of the issue *Passage Home* addresses. Moreover, they could even empathise with Mrs. Smith's actions, instead of recognising her behaviour as racist and wrong. Further research is needed to validate these findings (Olson & Harrell, 2020). Although *Passage Home* has a very strong focus on racial bias, and less on diversity, it similarly makes use of experiencing situations through the perspective of another as a learning tool. Therefore, it is relevant to *Diverse Perspectives*.

2.4 The role of diversity

This section elaborates on the concept of diversity and how it is linked to *Diverse Perspectives*. It discusses research on diversity interventions and how those results could be beneficial for the game.

2.4.1 What is diversity?

As mentioned before, diversity consists of the (in)visible differences that exist between people. It is a characteristic that belongs to a group, instead of an individual. This is because a group allows individuals to discover that others are different from themselves. Moreover, diversity should be seen as a whole and complex concept, since it comprises more than just the sum of several features. Diversity can be split up into two levels: 1) surface-level, which are the visible attributes, and 2) deep-level, which are the non-visible attributes (Gómez-Zará et al., 2020).

While race, ethnicity and gender are well-known aspects of diversity, it contains many more attributes than just those (Loes et al., 2018). A non-exhaustive overview of these (in)visible attributes and their different gradations is shown in Figure 2.5. It displays the *Academic Wheel of Privilege* (Elsherif et al., 2022). The wheel represents seven different categories that are divided into twenty aspects. For each aspect, a circle closer to the inner ring implicates

Figure 2.5: The Academic Wheel of Privilege. It shows a non-exhaustive overview of diversity presented in seven different categories that are divided into twenty aspects. The closer an individual is located to the inner ring, the more likely they can benefit from privileges (Elsherif et al., 2022).

a higher chance of having a certain privilege. Being in the inner ring on (almost) all aspects invokes the *mythical* norm. This norm usually consists of people who are "white, thin, male, young, heterosexual, Christian, and financially secure" (Lorde, 1984, p.115). On the contrary, when one's attributes are mostly located on the middle or outer rings, they are likely to have less privileges and could even be marginalised.

The interplay of multiple marginalised aspects invokes *intersectionality*. While this topic is considered to be out of scope for this thesis due to its complexity, we will briefly touch upon it. This is because diversity is associated with intersectionality. Intersectionality occurs when marginalised aspects *intersect* with each other. For instance, a person can both be Black and a woman. Here, these aspects interplay with each other and cause different lived experiences compared to a man who is Black or a woman who is White. That is, the Black woman has a chance of experiencing discrimination based on both her gender and her skin colour (Crenshaw, 1988; hooks, 1991).

To return to the topic of diversity, it should be taken into account that it is a 'double-edged sword'. On the one hand, diversity in teamwork could increase the information available in the group and therefore the group's performance, as people have different backgrounds. On the other hand, diversity could hinder teamwork, as people differ too much from each other to find common ground to work together effectively. Moreover, subgroups could form, with distrust and conflicts as result. Approaches to weaken these subgroups could consist of creating an overarching identity, stimulate collaboration and training in perspective taking (Homan, 2019).

2.4.2 How to encourage diversity?

2.4.2.1 Focus on behaviour directly instead of attitudes

Dwyer and Smith (2020) researched the effects of a single mandatory diversity and inclusive teaching workshop. The workshop durated 75 minutes and was given at the authors' university. Their results indicated the workshop had a positive impact, as its contents were remembered after more than one year. While mandatory workshops could invoke resistance, it is still valuable to conduct them, as it makes people familiar with topics of diversity and inclusion. These topics should not be seen in isolation, but rather be integrated in the social skills that are needed in daily life.

Yet, diversity trainings have mixed results and could even have the opposite effect, creating backlash and unintended consequences (Noon, 2018). For instance, Moreu et al. (2021) state that obligatory diversity training is often met with anger and resistance. Even some trainings resulted in an increased hostility towards marginalised groups afterward. Furthermore, Pietri et al. (2019) reported that when people's knowledge of bias increased and inequalities were addressed, it could do more harm than benefit to these groups, since it highlights their stigmatised position. In other words, being aware of biases is not automatically equal to behaviour change (Noon, 2018).

Homan (2019) states that "diversity mindsets", which are "mental representations of team diversity" (p.6), have a significant role in handling group diversity and interventions on that topic. These mindsets could be influenced by several factors, such as openness to experience, cultural intelligence and multicultural experiences. That is because such traits make people more open to different opinions and ideas and thus reduce their intergroup biases as they rely less on their automatic categorisations. Nevertheless, the author states that there is still a lack of information on effective interventions. Additionally, the evidence for a link between attitude change and behaviour change is weak, according to Moreu et al. (2021). The authors argue that it is possibly more effective to focus directly on behaviour that encourages diversity, rather than trying to change people's biassed minds. Thus, it seems more beneficial for *Diverse Perspectives* to focus on explicit behaviour guidelines, as these effects seem to have more impact than changing attitudes, which are implicit.

2.4.2.2 Three steps to consider in an intervention

As mentioned in section 1.1.1, the goal of *Diverse Perspectives* is to stimulate behaviour that promotes diversity. Moreu et al. (2021) describe three steps to consider in an intervention, which consists of: 1) selecting a target behaviour, 2) selecting a target audience, 3) explaining the 'barriers' and 'benefits'.

Step 1 means identifying a behaviour (or a small set of behaviours) that, once executed, has the biggest positive impact on the marginalised groups. In other words, what behaviour is desirable in promoting diversity? And what behaviour is not? Factors that play a role in choosing a behaviour are its impact, the success probability and its reachability. To determine what positive behaviours should be encouraged and what negative behaviours should be avoided, initial interviews were held with (international) students at UU, as they are the target users of the game. More information is present in section 3.3.1.

The second step consists of selecting the target audience that should adopt the target behaviour. Although the people that have the strongest biases about others might be the ones that could benefit the most from such intervention, it is suspected that the probability

for successful behaviour change is low, since it would be hard to challenge their deep-rooted beliefs and behaviours (Noon, 2018). Instead, it seems more fruitful to target the 'middle group' that does not have strong opinions about diversity, but is receptive to behaviour change and is willing to familiarise themselves further with the topic. In other words, (international) students at UU are perceived to be suited as a target audience.

The last step is exploring the barriers and benefits of the target behaviour, which is also an aim of the initial interviews. Barriers can contain anything that prevents one to execute the desirable behaviour. For example, it could be the case that members from this middle group avoid interactions with outgroups because they are afraid they would behave inappropriately due to a lack of knowledge or experience. Consequently, they rather not interact at all than run the risk of having a negative interaction (Harrison-Bernard et al., 2020). Alternatively, people may approach others with openness and curiosity, while experiencing discomfort and anxiety as well, since they perceive the other as unknown and different (Fuertes et al., 2000). To overcome this barrier, a good first step to this would be to familiarise these people to different and challenging situations in a safe environment, such as a workshop or a serious game. Benefits define the positive effects when the desired behaviour is shown. Examples could be more interaction between different groups and less intergroup conflicts.

2.5 Main findings of the Related Work

There are many related terms in the domain of serious games, such as *entertainment games*, which purely have entertainment as the main goal. In the case of *gamification*, game-like elements are added to a non-game. *Game-based learning* is a learning technnique that uses serious games as *product*. Although no single theory could explain how serious games work (Krath et al., 2021), we highlighted three: 1) SDT, 2) flow, and 3) cognitive load. Since serious games have an educational component as well, we elaborated on two pedagogical learning theories: ARCS and 8LEM.

Next, different types of narrative and progression are discussed. Carstensdottir et al. (2019) described six types of narratives: 1) linear, 2) branching, 3) foldback, 4) broom, 5) hidden, and 6) opportunistic. Additionally, the authors explained four types of progression: 1) choice, 2) scripted scenarios, 3) discovery, and 4) in-game systems. As *Diverse Perspectives* is a visual novel, the story plays a vital role. It uses a branching narrative that the player can progress through selecting a choice. Additionally, Øygardslia et al. (2020) illustrated six design principles for visual novels. *Diverse Perspectives* aimed to take these into account.

Moreover, related serious games in the domain of inclusion and diversity were discussed. Here, *Change Perspective* and the *Dilemma Game* served as key examples, as the combination of these two forms in essence the basis of *Diverse Perspectives*. Other serious games that were presented, are *MASELTOV*, *Secrets of the South* and *Passage Home*.

Lastly, the role of diversity was explained. Diversity is complex and multi-faceted, as the *Academic Wheel of Privilege* (Elsherif et al., 2022) depicts. To promote diversity, people should be at least familiarised with this concept. Yet, mandatory diversity trainings and workshops could even have the opposite effect (Moreu et al., 2021). Therefore, it seems to be more fruitful to focus directly on behaviour, instead of attitudes, to encourage diversity. Moreu et al. (2021) even describe a three-step plan for such an intervention. However, it is important to keep in mind that a change in attitude does not automatically lead to a change in behaviour (Noon, 2018).

Chapter 3

Design

This chapter describes the game mechanics that are chosen for the design of *Diverse Perspectives* and sketches out the content of the story. Furthermore, it discusses the steps of the design procedure that are held to create the game, including the initial interviews with target users.

3.1 Game mechanics

Game mechanics describe the actions players can do and how the latter sees and experiences the game world. Game mechanics are the interactions that remain when a game is stripped from all graphical representations, technology and narratives. Examples of game mechanics are story, characters, actions and rules (Mildner & Mueller, 2016; Schell, 2015). Since *Diverse Perspectives* builds further on *Change Perspective*, the former shares many similar game mechanics with the latter. The game mechanics for *Diverse Perspectives* are described below.

Diverse Perspectives is a singleplayer visual novel game. The reason why it is a singleplayer game, is because the player's choices affect the story of the game. In real life, players have to make choices on their own as well. The story of the game is narrative-driven and it progresses via a path over the multiple branches and the player can reach multiple endings. While multiple endings could confuse the player whether they reached the 'real' or the 'best' ending, they could encourage them as well to replay the game. An advantage of multiple branches and endings is that they give the player the feeling that they are in control. Since they are able to choose the narration themselves, it could increase their motivation (Naul & Liu, 2020). Moreover, freedom of choice fulfils the autonomy need of intrinsic motivation, as stated in the SDT (Deci & Ryan, 2008). More information about branches can be found in section 2.2, while the content of the story is elaborated in section 3.2.

Moreover, the game is displayed in multiple coloured 2D frames. Each frame could show the avatar, a NPC and their dialogue boxes. The layout and navigation are intended to be simple and clear, as we aim to reduce the ECL as much as possible. In the background, several environments are visible, such as places around the university campus and the city centre. These environments are chosen, because according to both the experiential learning theory and situated learning theory, contexts that are similar to the real-world could benefit problem-solving skills (Krath et al., 2021; Paletta et al., 2013). Furthermore, it aims to satisfy the relatedness component in the SDT, as the surroundings should be familiar to UU students

(Deci & Ryan, 2008). Moreover, the avatar is shown from Third-Person Perspective (3PP), as this is common in visual novels. Additionally, it is important that the player could identify with the role their avatar plays, as identification seems to be related to learning outcomes (Øygardslia et al., 2020).

The game works via a point-and-click system, so the player has to click on boxes or buttons to proceed. The game has no timing or random elements and its only rule is that once a choice is made, it is final. In other words, the player cannot go back in the story to change their choice.

3.2 Content of the story

The main story is based on the diversity dilemmas in the context of student inclusion. Concretely, this means that the choice options that are offered in the story are based on the *Dilemma Game*, instead of based on cultural questions, as was the case in *Change Perspective*. Each choice option can be presented by a different NPC and has its own (sub)theme that fits the NPC. For example, a choice option could be about how to use someone's pronouns if they are non-binary. The themes are based on the *Dilemma Game* and the *Academic Wheel of Privilege*. Another reason to base it on the *Dilemma Game* is because it forms a clear link to the diversity aspect. Likewise, the choices are focused on diversity to directly appeal to the player's behaviour, since their choices shape the outcome. Hence, the link between target behaviour and target outcome is intended to be explicit.

The side story is still about student inclusion. The scenario is that the player is a new (international) student at UU. They need to find a room in Utrecht and they can do so by building and increasing their social network. The bigger their network, the higher the chance of finding a room. This specific scenario is chosen, because finding a room is a relevant issue as a student. Therefore, it aims to make the situation of the game relatable by the player, regardless of their nationality. By choosing the right options, which is showing inclusive behaviour, the player can befriend the NPC and thus increase their social network.

In other words, *Diverse Perspectives* intends to include the educational message by allowing the avatar to participate in several conversations with NPCs in order to present new knowledge to the player. For example, the game has several moral dilemmas, as the choices the player makes will have major or minor consequences to the narrative. Such a dilemma could be morally right (good option), wrong (bad option) or neutral (ambiguous). Presenting the player different options provides them some sense of agency, which is essential to achieve player engagement.

Moreover, *Diverse Perspectives* uses 'hidden agency', where seemingly trivial choices could have an influence on the story later (Tancred et al., 2018). An example of this is the trivia question about the height of the Dom Tower. If the player chooses the correct answer, they unlock a bonus dilemma. This bonus dilemma provides the player an extra opportunity to change the outcome of the game.

3.2.1 Narrative structure

The narrative structure is drawn out in a graph, so that the different branches are clear. A part of this graph is shown in Appendix A.1, while the full Miro board can be seen here[1].

[1] https://miro.com/app/board/uXjVPihKNDs=/?share_link_id=296409787313

The story will contain crucial decisions that affect the outcome, and everyday decisions, which do not affect the outcome. The topic of crucial decisions are diversity matters, while the everyday decisions are small decisions, such as taking the bike or public transport to university. However, the player will not be explicitly informed about what type of decision each choice is. In the end, the choices should teach the player both about diversity and Dutch student life at UU.

The structure is chronological and the time span is one day. That means that the story starts with the avatar waking up and ends with the avatar going to sleep. Between the beginning and the end, the player goes through several scenes that could belong to the daily routine of a student. Examples are having breakfast, travelling to campus, following a lecture, spending their free time in the city centre, cooking dinner and going to a *hospiteeravond*. The reason why a chronological order is chosen, is to make the narrative relatable and realistic to students. Flashbacks or flash forwards are not integrated, as this is not possible in real life.

3.2.1.1 Version A: with perspective switch

In this version, the player can choose between a Dutch avatar or an international avatar. It does not matter which one they select, as this choice hardly influences the story. Throughout the game, a perspective switch could happen between the player's avatar and a NPC that they are interacting with. During this switch, the player jumps into the mind of the NPC. This is meant to explain the reasoning behind the NPC's behaviour when a dilemma occurs, so that the player receives the NPC's viewpoint as well. After reading the NPC's thoughts, the player returns to their own body and the switch is over.

It is important to know that the perspective switch only happens when the player is in proximity with the other character they could have chosen as avatar. That is, if they play as a Dutch avatar, the other character represents the international student and vice versa. Because of the way the narrative is structured, the perspective switch occurs at least once and at most seven times, but the total number of switches depends if the player befriends the other student and tags along with them or not. In other words, befriending the other student is 'rewarded' with perspective switches.

3.2.1.2 Version B: without perspective switch

This version is identical to version A, except there is no perspective switch. This also means that the playtime of this game is slightly shorter, as there is no jumping between characters, which results in less scenes.

3.2.2 Take-away messages at the end of the game

The game allows the player to practise several diversity scenarios in a safe environment. To make the educational content more explicit and clear, *Diverse Perspectives* gives the player some take-away messages once they complete the game. These messages consist of a recap of the (sub)themes they could have encountered in the game and a small "Do & Don't". These take-aways are meant to serve as practical guidelines that players could use in the future.

Beside the explicit take-aways, the game intends to 1) teach the player a little about Utrecht and its student life, 2) encourage interactions with others that differ from oneself, and 3) show that small actions could have big consequences.

3.3 High level overview of the iterative design procedure

Since *Diverse Perspectives* is a serious game with educational content, it should be carefully designed. This usually happens in an iterative process, which involves the steps of a literature review, creating a prototype, iterative testing to refine the prototype, evaluation of the prototype, and academic documentation (Øygardslia et al., 2020). Table 3.1 shows a high level overview of the iterative steps, its purpose, how participants are sampled and how many participants are aimed for. Note that each step that includes participants had a pilot test beforehand to test the setup.

For the whole iterative design procedure, the Ethics and Privacy Quick Scan of the Utrecht University Research Institute of Information and Computing Sciences has been conducted. It classified this research as low-risk with no fuller ethics review or privacy assessment needed. The results are documented in a separate file called '`Ethics and Privacy Quick Scan.pdf`' and can be shared upon request.

Table 3.1: The steps of the iterative design for *Diverse Perspectives*. Each step is accompanied with its purpose, participant pool and aimed number of participants.

Step	**Purpose**	**Participants sampling**	**Aimed number**
Initial interviews (section 3.3.1)	Inspiration for scenarios and choices in the game	Convenience and snowball	5
Low-fidelity prototype (Figma)	To be used in the focus group	N/A	N/A
Focus group to evaluate the low-fidelity prototype	Gather feedback on usability, fun, aesthetics, User Experience (UX), etc.	Convenience and snowball	5-10
High-fidelity prototype (Ren'Py)	To be used in the experiment	N/A	N/A
Experiment to evaluate the high-fidelity prototype (section 5.3)	Answers to the research questions	Convenience and snowball (*INFOMSEGA* and *INFOB3APGA* students, colloquium attendees)	30

3.3.1 Initial interviews

As mentioned before, initial semi-structured interviews were held with the target audience to form concrete ideas on what the target behaviour should be. Involving the target users in the early design stages is perceived to be beneficial for the success of the game (Paletta et al., 2013). As the content of *Diverse Perspectives* is based on the *Dilemma Game*, choice

examples from this game were used to shape the target behaviour. Supplementary material can be found in Appendix C and H (Blandford et al., 2016).

In total, five Human-Computer Interaction (HCI) students were interviewed (2 male, 3 female), aged between 22 and 24 years old ($M = 23.3$, $sd = 0.8$). Three participants had Dutch nationality, while one participant was Irish and another participant was German. Four participants identified as White, and one participant identified as Asian. The time each participant studied at UU varied between 0.5 and 5.5 years ($M = 2.3$, $sd = 2.1$). Every participant participated voluntarily and signed the informed consent form. Besides offering the participants a small refreshment and thanking them for their time, there was no further compensation for their participation.

Here, we will describe a summary of the participants' answers. All of them have heard of serious games before and some participants even mentioned examples. Yet, the *Dilemma Game* was only known to a few participants. However, those participants thought it represented the *Dilemmas of the Scientist*, which is a similar game they played in another course. In other words, no participant was familiar with the *Dilemma Game* that was described in section 2.3.3. After the researcher showed the participants an example dilemma from the latter game, participants were asked what they thought the purpose of the game was. Their answers varied from raising awareness and consciousness for marginalised people, creating safe spaces, promoting diversity and inclusion, and encouraging behaviour change. Moreover, each participant has heard of the terms 'diversity' and 'inclusion' before, and could roughly explain in their own words what the concepts meant to them.

When asked for example dilemmas, seven unique scenarios were mentioned in total. First, a participant described a situation where a student's non-binary pronouns conflicted with the religion of a teacher. Another situation involved people not respecting someone's pronouns just because they did not seem to like that person and they were not present in that conversation. A different topic that was mentioned a few times were language and culture barriers. Regarding language barriers, participants described situations where Dutch people refused to speak English to non-Dutch speakers. Likewise, a Dutch participant stated they found it difficult to both include themselves and others when choosing which language to speak.

On the topic of culture barriers, a participant stated that some customs could be considered as appropriate in certain cultures, while those same behaviours could be seen as offensive in other cultures. For example, eating a particular kind of meat. Additionally, several participants noted that Dutch and international students are being treated as separate groups. For instance, they mentioned that some introduction activities were specifically tailored to international students. Another example concerned feeling 'othered' in a conversation due to one's nationality. Hence, these scenarios were worked out in further detail in the narrative of *Diverse Perspectives*, as they are deemed to be relevant.

Furthermore, when participants were asked for topics that should certainly be represented in the game, their replies consisted of: 1) cultural awareness, 2) covert racism that is downplayed as 'just a joke', 3) internationalisation, 4) first-generation students or people who finished another educational level first, such as HBO (i.e. University of Applied Sciences), 5) awareness for LGBTQ+-members, people of colour, differences in religions and ages. In other words, a great variety of topics was mentioned and therefore *Diverse Perspectives* aimed to include at least these matters.

In sum, the examples and feedback participants provided is perceived to be very valuable and useful for *Diverse Perspectives*.

Chapter 4

Implementation

This section describes the implementation and evaluation of the low-fidelity prototype and the implementation of the high-fidelity prototype.

4.1 Low-fidelity prototype

The low-fidelity prototype is made in Figma[1]. The prototype consisted of many individual frames. Each step it took for the story to progress, a new frame was shown. To make the prototype interactive, these frames were linked together. Clicking certain buttons and arrows would trigger the appearance of the next frame. However, due to the nature of Figma being a drag-and-drop program, there were some limitations. Because conditional logic was not possible to implement, each consequential choice that would trigger a new branch had to be copied from the main branch. Concretely, this means that there were already 2 x 2 x 3 = 12 branches (Dutch vs International, Bike vs Bus and Together vs Alone-Together vs Alone).

Additionally, Figma is not able to trace the player's location on the branch, so some frames were skipped as they only made sense if a certain frame was shown before. Likewise, this also means that there was no keeping track of a score, so there was less fine-grained control over the story flow than intended. That is, once the player made a bad choice in the beginning, they were 'stuck' on that path to receive the bad ending, regardless of how many good choices they made afterwards.

A frame could consist of several elements: a character, a dialogue box, a choice option and a background. There were several criteria kept in mind when selecting characters: they should fit the scenario they represented, they should pass as a student and they should be diverse. Yet, the characters were still quite one-dimensional, as they were all good-looking, a bit on the older side and very stock-photo-like. To mitigate this, the characters could be edited to look more in-game like by using visual filters, for instance. Alternatively, real-life students at UU could pose as photo models. Nevertheless, these options were not used, as they served their function for the low-fidelity prototype.

When selecting the backgrounds, it was important that the picture resembled the scenario. For example, if a dilemma occurs in a lecture room, then the background should be a lecture room. Moreover, the background pictures should not be too outdated. This is to make the game a little bit more realistic. As the story takes place in Utrecht and on campus,

[1] Low-fi prototype

backgrounds that displayed these surroundings were chosen. All backgrounds were edited to fit the dimensions of the screen and a blurred copy was made. This blurred copy was used when characters were superimposed on the background, which is visible when they are having a dialogue. The blurring is done to help players focus on the text boxes and alleviate information overload.

4.1.1 Evaluation of the low-fidelity prototype

Two focus groups were held for evaluation. Supplementary material can be found in Appendix D and H (Blandford et al., 2016). In total, seven HCI students participated (2 male, 5 female), aged between 23 and 26 years old ($M = 23.6$, $sd = 1.1$). Four participants had Dutch nationality, while one participant was Cypriot, one participant had US nationality and another participant was German. Five participants identified as White, one participant identified as South Asian and another participant was bi-racial. The time each participant studied at UU varied between 0.1 and 5 years ($M = 1.8$, $sd = 2.0$). Each participant participated voluntarily and signed the informed consent form. Besides offering the participants a small refreshment and thanking them for their time, there was no further compensation for their participation.

There was a week between the first and the second focus group. This time interval was used to incorporate some feedback from the first focus group before the second focus group occurred. Specifically, the moral pace of the story was lowered by adding more trivial questions and elaborating the conversations on some dilemmas. Likewise, some scripts were rephrased to sound more natural. Moreover, more context was given to clarify the relevance of certain dilemmas. For instance, the reasoning behind the NPC's behaviour would be explained, to help the player understand their side of the story. Lastly, back buttons were added when big pieces of information were given, to allow the player to navigate back if they wanted to re-read something. We refrained from implementing back buttons before, since we wanted to prevent players from changing their decisions.

Feedback from the second focus group involved suggestions of indicating more clearly when the player reads the NPC's mind. For instance, this could be by giving the dialogue box another colour or adding a foggy filter. Allowing players to go back to the question was another piece of advice. Another suggestion would be to extend the game by introducing the player to Utrecht more, just like in introduction week, as well as introducing the concept of the perspective switch when this occurs for the first time. Furthermore, the choices could be more nuanced, as the participants indicated that the choices they made in the game would differ with what they would do in reality. They suspected what the 'right' answer was, but in real-life they would not confront strangers.

The general setup of pre-test questionnaire, game and post-test questionnaire worked well. The questionnaires were perceived as understandable, sensible and well-balanced. A detailed description of them can be found in section 5.2. Participants liked that the game was very Utrecht focused, had different choice options and showed scenarios that one may not encounter everyday, but were still realistic. Likewise, the diversity of topics and the recap at the end were appreciated and the perspective switch added an interesting angle to a dilemma. Moreover, the layout was intuitive and navigation in the game was easy. While some participants preferred a 1PP, other participants liked the face-to-face communication that is displayed in 3PP, so it was clear who was talking.

In sum, *Diverse Perspectives* should be extended, as participants deemed the game rather too short instead of too long. Concretely, the narrative of the high-fidelity prototype should be more nuanced, so that the player's choices would be more realistic. Otherwise, the possibility exists that players would feel pressured to pick the 'best' option or they felt 'graded'. Moreover, there should be more variation in the avatars to allow player identification. Lastly, more context should be added to the game, such as introductions to the city, the characters and the perspective switch, to increase player engagement in the story.

4.2 High-fidelity prototype

The high-fidelity prototype is made in Ren'Py[2] version 8.0.3. This is a free game engine to create visual novels in Python. PyCharm version 2022.1 was used as IDE. Various materials were reused from the low-fidelity prototype, such as the storyline and the layout. For example, the background images were re-used. Moreover, feedback from both focus groups was incorporated in the prototype. For instance, the storyline was extended a bit, as more trivial choices and trivia facts were added. This is done to lower the moral pace of the game and add some fun and light-hearted elements. Additionally, the dilemmas were nuanced a bit, to appear more realistic. That is, extreme choices are left out, as it is unlikely anybody would behave in such a manner in real life. More dilemmas have follow-up choices as well, so the player receives a 'second chance' to let them reflect on a dilemma a bit more. Moreover, the text font is displayed in italic to clarify to the player that they are reading the thoughts of a character. Furthermore, the first perspective switch is introduced a bit more.

Additionally, animations were added, such as transitioning to a new background when the scene switches and characters that fade in and out when they speak. Likewise, a functional score system is in place to determine which ending the player should receive. The complete narrative can be found in Appendix H. All game files and additional materials can be found on my GitHub[3]. However, the word plays on existing places such as café *BodyWalk*, study association *Tricky* and supermarket *Jumborama* have been changed to their original name. Respectively, these are *BodyTalk*, *Sticky* and *Jumbo*. This is done to avoid confusion when mixing facts and fiction, as more real educational content about Utrecht is added to the high-fidelity prototype. This Utrecht content is taken from Nieubuur (2020, 2022) and the researcher's own knowledge.

4.2.1 Story

As stated in section 3.2, the time span of the whole story is a day. The player steps into the shoes of an UU student. The summer break has just ended and they are about to begin their very first day at Utrecht University. Moreover, they do not know anyone yet, since they are new in Utrecht. Because their hometown is far away from the university, they are desperately looking for a room in Utrecht. Currently, they are staying in a hostel room. This evening, they are invited to a *hospiteeravond* (*hospi*) in a student house called *Huize Peereboom*. Moreover, the house has space for two new roommates! However, the catch is that it is a mixed house, so they are looking for one Dutch roommate and one international roommate. This is because the house wants to maintain the balance between different nationalities.

[2]https://www.renpy.org/
[3]https://github.com/geertjepetersrit/master-thesis

Likewise, *Huize Peereboom* finds it important that everyone gets along well, so they prefer Dutch and international students who are already close. In other words, it is in the player's benefit to befriend someone from a different nationality and to interact with strangers as it could increase their chances of getting a room. One may never know when you run into your potential roommates...

In total, nine (sub)topics of diversity dilemmas were presented in the story. These are: 1) cultural differences: ethnicity, 2) identity: transgender and non-binary people, 3) disabilities: physical, 4) identity: sexual orientation, 5) formal education, 6) neurodivergence, 7) cultural differences: language, 8) cultural differences: religion, and 9) dietary wishes.

4.2.2 Endings

The player's choices affect the ending. There are three possible endings: good, neutral and bad. However, the goal of the game is to reach the good ending. In the good ending, both the player and their new (international) friend get the room. In the neutral ending, neither of them gets the room. In the bad ending, the player does not get the room, while the other person does.

To get the good ending, the idea is that more interaction leads to a greater network, a better first impression at the *hospi* and hence a bigger chance of being picked as the next roommate at *Huize Peereboom*. More interaction also means thinking and reflecting about the educational content and being able to make the right decisions.

4.2.3 Choices

There are two types of choices. Consequential choices have a major effect on the outcome of the story, while trivial choices do not.

4.2.3.1 Consequential choices

The topics of these choices are based on the *Dilemma Game* and the initial interviews. The perspective switch only occurs after each consequential choice that involves the other person. An exception holds for Dilemma #8A, Dilemma #9 and Dilemma #10, as they are not followed by a perspective switch. This is to limit the amount of switching, as it could be otherwise too complex to follow. Further details about the switch were explained in section 3.2.1.1.

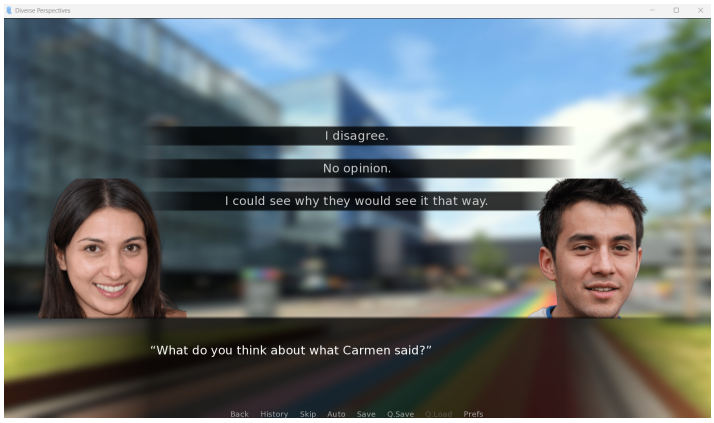
At every choice, the words in **bold** (see Appendix I) serve as keywords, which are shown as choice options to the player. The sentences that follow consist of the full response. A choice can be good, neutral, or bad, assigning +1, 0, or -1 points respectively. However, note that a situation does not need to exactly contain these three choice categories. Moreover, the different choice categories are presented in a random order. This is to prevent learning effects.

If the player chooses the 'good' option in Dilemma #4, the score will be doubled, instead of incrementing it with 1. This is because it involves a NPC who is a member of *Huize Peereboom*. Therefore, the highest score the player can get is 13 points and the lowest score is -5 points.

Note that the situations and choice options are a simplification of reality. The 'right answer' is not always clear in real life, as the lines between good, neutral, bad and everything in between are often more blurry and nuanced.

An example of a consequential choice is depicted in Figure 4.1. It shows Dilemma #3A, where the player encounters someone who refuses to use non-binary pronouns (Figure 4.1a). In Figure 4.1b, the player can choose one of the three options as a response.

(a)



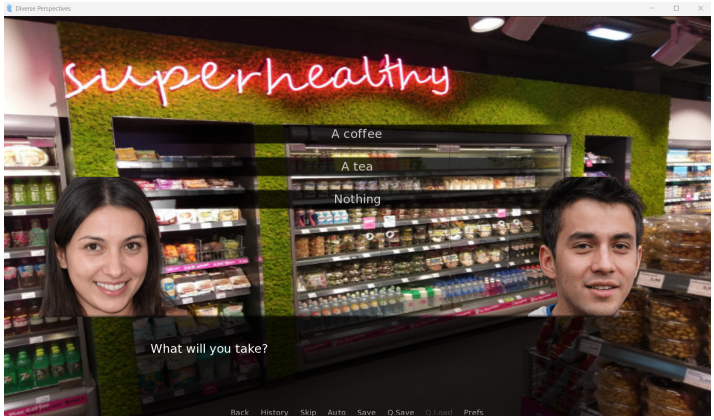
(b)

Figure 4.1: Screenshots of Dilemma #3A in *Diverse Perspectives*. The player encounters a situation (a) and is then offered to choose a response option (b).

4.2.3.2 Trivial choices

These are everyday decisions in Dutch student life. The examples are based on *Change Perspective* and the researcher's own experiences. Topics of these choices include avatar customisation, eating breakfast, travelling to campus, having a drink, answering some trivia and cooking dinner. Figure 4.2 illustrates two trivial choices. Figure 4.2a allows the player to either snooze their alarm or to get up. In Figure 4.2b, the player can pick a drink in the *Spar University*.

(a)



(b)

Figure 4.2: Screenshots of trivial choices in *Diverse Perspectives*. The player can choose what to do after their alarm went off (a) and what drink they would like in the *Spar University* (b).

4.2.4 Avatar sprites and behaviour

The player can customise a few aspects of their avatar, such as their gender and their nationality, but the sprite and the personality is pre-defined. These aspects should not be overlooked, as the avatar's behaviour could translate to the outside the game world. Examples are that players could adapt to certain clothing, hairstyles or even attitudes that are similar to the avatar. It is also possible that this effect occurs the other way around. That is, the avatar's behaviour is influenced by the outside world. This is termed the *Proteus effect*. The extent to which the Proteus effect happens, depends on the game mechanics, such as degree of avatar customisation, difficulty of the game, graphical representation and controller type (Szolin et al., 2022, 2023).

Two theories attempt to explain how the Proteus effect could function. First, the priming theory states that the avatar serves as an external cue that triggers schemas about a concept. In other words, the avatar invokes a certain set of behaviours that is associated with them. The second theory, self-perception theory, posits that beliefs and attitudes follow from self-observed behaviour and are used to explain an individual's actions. Since the avatar is seen as

an extension and embodiment of the player, the player will align their behaviour and attitudes to match that of their avatar. Nevertheless, the two theories could also be interconnected (Szolin et al., 2022).

In total, there are three avatar sprites: female, male and non-binary. In the end, avatars of non-existing people were generated with AI via this website[4]. Once downloaded, the watermark was cropped off, as it should not be visible in the game. Then, the images were made transparent and resized to fit into the game. In total, 12 characters were generated, where three of them could be chosen as the avatar (female, male and non-binary). One criteria that was kept in mind was that the characters should be diverse, both on surface-level and deep-level. On surface-level, different genders and skin colours were generated, while on deep-level, each character presented their own dilemma. This is because representation does matter, especially for people who identify with multiple marginalised groups (Smith & Decker, 2016). Therefore, both the avatars as well as the NPCs are intended to represent a diverse group of students. In other words, the masculine, white, and heterosexual gaze should be avoided (Tompkins & Martins, 2022).

The reason to choose for non-existing characters is to avoid placing photos of existing people in situations that can be interpreted as inappropriate. While stock photos were used in the low-fidelity prototype, these are still real people and the game does not aim to portray that specific person as offensive. Therefore, the generated avatars served as a workaround for this problem.

[4]https://this-person-does-not-exist.com/en

Chapter 5

Method

This section describes the setup of the experiment, which materials and measurements were used, which variables were included and how the procedure was executed.

5.1 Participants

The target group for *Diverse Perspectives* consists of students at UU. Therefore, participant recruitment happened within this target group. It occurred through convenience and snowball sampling. Moreover, two short pitches were held in the courses *Serious Gaming (INFOMSEGA)* and *Gamification and Applied Games (INFOB3APGA)* and a colloquium presentation was held to promote the research.

In total, 32 participants were recruited (8 male, 24 female). There were 20 Dutch students and 12 international students. Moreover, 24 participants identified as White, while 8 participants identified as non-White. In addition, 10 participants identified themselves as part of a marginalised group, while 19 participants did not. 3 participants did not know or preferred not to say. Furthermore, 9 participants experienced discrimination or exclusion based on their own aspects, while 18 participants did not. 5 participants did not know or preferred not to say.

The rest of the participant demographics can be found in Table 5.1. A 5-point Likert scale was used to measure familiarity, where 1 = not at all familiar and 5 = extremely familiar.

Table 5.1: The demographics of the participants (N = 32) who participated in the experiment.

	Mean	**Standard deviation**
Age	23.0	3.1
Years at UU	2.3	1.7
Familiarity with diversity and inclusion	4.0	0.8
Familiarity with serious games	3.3	1.3

Each participant participated voluntarily and signed the informed consent form. As compensation for their participation, participants were offered refreshments in part 1 of the ex-

periment. Upon completing part 2, they were rewarded with a copy of the game, their personal player statistics and having a chance in the raffle to receive a special mention in the Acknowledgements.

As the experiment had a between-subjects design to avoid learning effects, the participants were divided into two conditions: the experimental condition *with* the perspective switch (version A) and the control condition *without* the perspective switch (version B). Each participant received a unique participant ID. If this number was odd, they were assigned to the experimental group. If their number was even, they participated in the control group. Therefore, 16 participants were put in the experimental group and 16 participants were allocated to the control group.

5.2 Materials and measurements

R1 states: *What is the impact of the perspective switch on one's attitude towards diversity?* This subquestion will be answered by measuring one's attitude towards diversity. There are several questionnaires to measure one's attitude towards diversity. While the Openness to Diversity and Challenge (Loes et al., 2018) has been considered, it was deemed too short and too straightforward for the purpose of the study. Instead, we decided to use the Shortened Version of the Miville-Guzman Universality-Diversity Scale. This scale measures one's awareness and acceptance of differences among people. It is validated and its Cronbach's α = .77, which is acceptable. The M-GUDS-S is divided into three subscales, namely *Diversity of Contact (DC)*, *Relativistic Appreciation (RA)* and *Comfort with Differences (CD)*. Each subscale has three items, thus in total the questionnaire contains fifteen questions. Each question is measured on a 6-point Likert scale, where 1 = strongly disagree and 6 = strongly agree.

The scores should be interpreted as follows: a higher score for DC indicates a higher frequency of contact with different cultures. Similarly, a higher score for RA suggests a higher appreciation of similarities and differences. Lastly, a higher score for CD implies a higher sense of *discomfort* when confronted with differences, hence this score is inverted (Fuertes et al., 2000; Kottke, 2011; Trolian & Parker III, 2022). The M-GUDS-S questionnaire can be found in Appendix F.

To measure as accurately as possible the difference in the player's attitude, the player filled in the M-GUDS-S questionnaire before playing the game to obtain their baseline (pre-test) and filled in the questionnaire again after playing the game (post-test). The time between the pre-test and the post-test was approximately a week. Thus, the difference in the player's attitude is the scores from the post-test subtracted by the scores from the pre-test. We hypothesise a positive difference in DC or RA, or a negative difference in CD, as this suggests a positive influence from the game on one's diversity attitude. Moreover, we speculate that these differences will be bigger in version A, due to the effect of the perspective switch. To answer R1, the hypothesis is split up in two parts:

- The alternative hypothesis of the first part $H1.1_A$ is defined as: *Both the scores for Diversity of Contact and Relativistic Appreciation are higher in the post-test, while the score for Comfort with Differences is lower in the post-test.*

- The null hypothesis $H1.1_0$ is defined as: *There are no differences in the scores for Diversity of Contact, Relativistic Appreciation and Comfort with Differences between the post-test and pre-test.*
- The alternative hypothesis of the second part $H1.2_A$ reads: *The differences between the scores on all three components of the M-GUDS-S are higher in version A than in version B.*
- The null hypothesis $H1.2_0$ reads: *The differences between the scores on all three components of the M-GUDS-S between version A and version B are equal.*

R2 is defined as: *Which factors contribute to a positive gaming experience?* And R3 poses the following: *What is the effect of identifying with a marginalised group on the gaming experience?* These subquestions will be answered by measuring the gaming experience. Two common scales to measure the gaming experience are the Player Experience of Need Satisfaction (PENS) and the Game Experience Questionnaire (GEQ). Johnson et al. (2018) validated them and reported that the GEQ is partially supported, while the PENS is largely supported. While the latter is more supported, it is deemed as less suited as well. This is because its dimensions apply to a lesser degree to *Diverse Perspectives*. That is, the PENS has its roots in SDT. Therefore, the PENS is especially suitable to measure motivation, identity and well-being. Since these concepts are not the core focus of *Diverse Perspectives*, the broader GEQ is presumed more applicable for evaluation.

Therefore, GEQ is chosen as a validated questionnaire to measure the gaming experience. It has a Cronbach's $\alpha = .81$, which is sufficient (Poels et al., 2007). Although the GEQ consists of three modules, only the first one is used, as it is the most applicable to the research. This module evaluates game experience on seven components: 1) Immersion, 2) Flow, 3) Competence, 4) Positive Affect, 5) Negative Affect, 6) Tension, and 7) Challenge. The GEQ contains 33 questions and uses a 5-point Likert scale, where 0 = not at all and 4 = extremely (IJsselsteijn et al., 2013). Likewise, only a subset of these seven components is used. This subset consists of 1) Immersion, 2) Flow, 3) Competence, and 4) Positive Affect. The reason why the first four components are used, is because they seem to be empirical supported to some degree (Johnson et al., 2018). The subset of the GEQ questionnaire can be found in Appendix G. Hence, we hypothesise that version A scores higher on all these four components than version B. This is because we suspect that the perspective switch adds an extra dimension to the situation, which is interpreted as positive. Thus, formally, our hypotheses are:

- The alternative hypothesis $H2_A$ is presented as: *Version A of the game scores higher on all four components of GEQ than version B.*
- The null hypothesis $H2_0$ is presented as: *There are no differences between the scores on all four components of the GEQ between version A and version B.*

Furthermore, we speculate that players who identify with a marginalised group score higher on all four components than players who do not, since they might perceive the game as more relevant to their own situation. Therefore, our hypotheses read:

- The alternative hypothesis $H3_A$ is represents the following: *Players who identify with a marginalised group score higher on all four components of GEQ compared to players who do not.*

- The null hypothesis $H3_0$ represents the following: *There are no differences between the scores on all four components of the GEQ between players who identify with a marginalised group and players who do not.*

5.2.1 Independent variables

There is only one independent variable, which is the version of the game. This independent variable has two levels: version A (with perspective switch) versus version B (without perspective switch).

5.2.2 Dependent variables

There are four dependent variables, which are: 1) differences in the results of the M-GUDS-S pre-test and post-test, 2) the results of the GEQ questionnaire on the first four components, 3) the consequential choices the player makes, and 4) which ending the player receives. The last dependent variable has three levels: good, neutral and bad.

5.2.3 Controlling variables

There are various controlling variables present. These are: age, gender, nationality, race/ethnicity, identification with a marginalised group, number of years studying at UU, familiarity with diversity and inclusion, and familiarity with serious games. The last two controlling variables are measured on a 5 point-Likert scale, where 1 = not at all familiar and 5 = extremely familiar.

While other aspects of diversity could be of interest, such as sexuality and (physical) disabilities, we decided not to ask for these. This is because the participant pool is not large enough to avoid individual identification once these variables are collected and from a single participant.

5.3 Procedure

The procedure consisted of a mixed-methods design. The goal of the experiment is three-fold. First, we aim to investigate the influence of the perspective switch in *Diverse Perspectives* on the player's attitude towards diversity. To assess the first goal, quantitative feedback is obtained via the M-GUDS-S questionnaire. Second, we wish to research the influence of this switch on the gaming experience. This goal is evaluated by gathering quantitative feedback from the four components of the GEQ. Third, we want to investigate the effect of identifying with a marginalised group on the gaming experience. This goal is assessed by using the GEQ in combination with the demographic information. Qualitative feedback in the form of semi-structured interviews after the game are held to complement the assessment of the goals. Supplementary materials can be found in Appendix E and H.

Therefore, the experiment is split in two parts: part 1 included the pre-test M-GUDS-S questionnaire, playing the game, the GEQ questionnaire and the semi-structured interview, while part 2 consisted of only the post-test M-GUDS-S questionnaire and the collection of demographic information. A schematic overview of part 1 and part 2 of the experiment can be found in Table 5.2 and Table 5.3, respectively. Part 1 was entirely performed on the researcher's laptop, which was an `HP Pavilion 13-AN1430ND`.

After filling in the pre-test M-GUDS-S questionnaire, the participant should call the researcher. Then, the researcher started up the game and explained the controls. After the participant finished the game, the participant should notify the researcher again, so that the latter could open the GEQ survey. When the participant finished the survey, they handed the laptop back to the researcher. They were asked if they would like a short break before conducting the interview. Lastly, part two was sent approximately a week later via email.

Table 5.2: Timeline of part 1 of the experiment. This part was done in-person. The total estimated time of the timeline is 45 min.

Task	**Materials**	**Time (min)**
Welcome and thank participants, explain the experiment and the steps	Information sheet and informed consent form, participant ID	5
Pre-test M-GUDS-S questionnaire	Qualtrics	5
Play the game	High-fidelity prototype of the game	10
GEQ questionnaire	Qualtrics	5
Semi-structured interview to gather qualitative feedback	Interview protocol and recording device	15
Thank participants and wrap up	A small reward, e.g. a snack	5

Table 5.3: Timeline of part 2 of the experiment. This part was done online. The total estimated time of the timeline is 5-10 min.

Task	**Materials**	**Time (min)**
Thank participants again, explanation of the follow-up	Information sheet and informed consent form, participant ID	2-3
Post-test M-GUDS-S questionnaire	Qualtrics	2-3
Demographic questionnaire	Qualtrics	2-3

Chapter 6

Results

This section discusses the analysis from the quantitative and the qualitative data. For the quantitative data, it describes the results of the hypotheses testing and the exploratory research. Regarding the qualitative data, it elaborates on the combination of a priori and emergent coding, and the formation of preliminary theories.

6.1 Quantitative analysis

The survey data was downloaded from Qualtrics and then pre-processed in Microsoft Excel. This means that unnecessary information, such as metadata in the columns and incomplete entries, were removed. Then, the Likert scales that were used in the M-GUDS-S and the GEQ were converted to numbers. As stated before, the M-GUDS-S used a 6-point Likert scale, where 1 = strongly disagree and 6 = strongly agree (Fuertes et al., 2000; Kottke, 2011; Trolian & Parker III, 2022). As each component consisted of 5 items, the maximum score for every component is 6 x 5 = 30 points. To reiterate, a higher score for DC and RA and a lower score for CD indicates a more open-minded attitude for diversity.

The GEQ used a 5-point Likert scale, where 0 = not at all and 4 = extremely (IJsselsteijn et al., 2013). Only the Immersion component consisted of 6 items, so that means that the maximum score for it is 6 x 4 = 24 points. The other three components each had 5 items, so that maximum score for every of them is 5 x 4 = 20 points. A higher score for a component suggests a more positive gaming experience.

Table 6.1 shows an overview of the analyses that have been performed per questionnaire. Note that for some tests, no hypotheses were defined beforehand. These are marked 'exploratory', as they fall under the exploratory research part. The other tests are described in the hypothesis part, which is the next section.

6.1.1 Hypotheses

After pre-processing, the data was analysed in R version 4.1.2. Each hypothesis followed the same approach. First, the assumptions were checked to decide which statistical test to use. Since the Likert scale was converted to numbers, it was categorical data. Although the numbers had an equal distance between them, the scales only consisted of 6 and 5 points, and therefore the data was treated at ordinal level (Joshi et al., 2015). That is, the normality

Table 6.1: Overview of the analyses that have been conducted per questionnaire.

M-GUDS-S	**GEQ**
Pre-test vs post-test	-
A vs B	A vs B
Marginalised (exploratory)	Marginalised
Gender (exploratory)	Gender (exploratory)
Nationality (exploratory)	Nationality (exploratory)

checks are left out from the descriptive statistics, since it violates that assumption. Furthermore, non-parametric tests were used. However, this means that the inferences from these tests are weaker than parametric tests, thus the results should be interpreted with caution (Field et al., 2012).

All hypotheses were tested with significance level $\alpha = .05$. Moreover, all box plots indicated some outliers. However, these were not removed, as they were not seen as errors, but rather as providing new insights to the data. For example, it could be that the outliers on the different components of the questionnaires belong to the same participant. Then these outliers could indicate that this particular participant tends to have a different opinion compared to the rest of the group. This suggests the potential of personalisation in serious games, as every player has their own preferences.

For conciseness, all Tables elaborating on the descriptive statistics can be found in Appendix B. The box plots, that visualise the same data, are present in the current chapter.

6.1.1.1 Differences between M-GUDS-S pre-test and post-test

Figure 6.1 displays the distributions of the three components of the M-GUDS-S in a box plot. The specific component and to what test it belonged can be seen on the x-axis, while the y-axis represents the score. It can be seen that the pre-test (blue) and post-test (yellow) scores do not differ much from each other for each component. Likewise, the scores for DC and RA are very similar, as a higher score indicates a more open attitude towards diversity. The scores for CD are in the lower region, as that scale is inverted. That is, lower scores indicate a more open-minded diversity attitude.

As stated before, the alternative hypothesis $H1.1_A$ is defined as: *Both the scores for Diversity of Contact and Relativistic Appreciation are higher in the post-test, while the score for Comfort with Differences is lower in the post-test.*

The null hypothesis $H1.1_0$ is defined as: *There are no differences in the scores for Diversity of Contact, Relativistic Appreciation and Comfort with Differences between the post-test and pre-test.*

Therefore, the Wilcoxon signed-rank test is used to test H1.1 for each component, as the samples are dependent. The results per component are as follows:

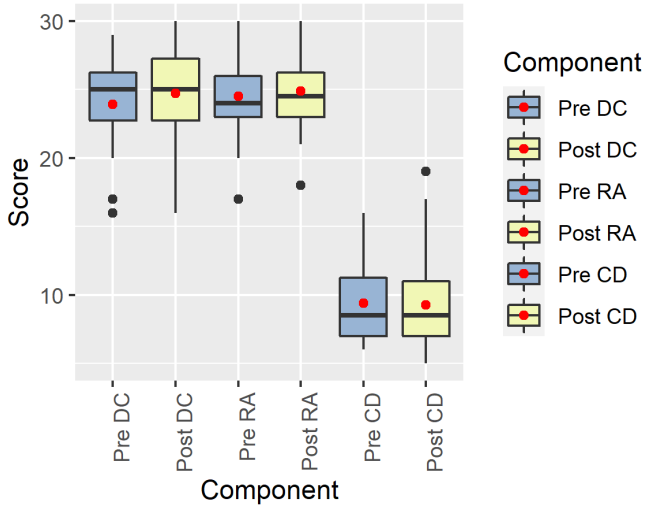


Figure 6.1: The distributions of the pre-test (blue) and post-test (yellow) M-GUDS-S data in a box plot. The three components (DC, RA and CD) are on the x-axis. The score of each component is on the y-axis.

- Post-test DC scored higher (median = 25.0, IQR = 4.5) than pre-test DC (median = 25.0, IQR = 3.5). According to a Wilcoxon signed-rank test (one-tailed) for dependent samples, this difference was significant, $V = 236$, $p = .02$. Furthermore, it represented a medium effect, r = -.28.
- Post-test RA scored higher (median = 24.0, IQR = 3.3) than pre-test RA (median = 24.0, IQR = 3.0). According to a Wilcoxon signed-rank test (one-tailed) for dependent samples, this difference was *not* significant, $V = 261$, $p = .17$.
- Post-test CD scored lower (median = 8.5 IQR = 4.3) than pre-test CD (median = 8.5, IQR = 4.0). According to a Wilcoxon signed-rank test (one-tailed) for dependent samples, this difference was *not* significant, $V = 143$, $p = .56$.

Hence, only $H1.1_A$ for DC can be accepted as its post-test score was significantly higher. However, this was not the case for RA and CD, so $H1.1_A$ is rejected.

6.1.1.2 Differences between versions in M-GUDS-S

To analyse whether there have been any differences between the pre-test and post-test scores between the two conditions, the dataset has been split into two: version A (the experimental group) and version B (the control group).

Figure 6.2 displays the distributions of the differences of each of the three components of the M-GUDS-S in a box plot, where version A (blue) is compared to version B (yellow). This can be seen on the x-axis. On the y-axis, the score difference is represented. Although results across the groups do not differ much from each other, the bigger IQR and whiskers for the experimental group indicate that their answers varied more than in the control group. Moreover, the mean and the median in each box plot lie close to zero. This suggests that the participants' answers to the M-GUDS-S did not vary much after playing the game, in general.

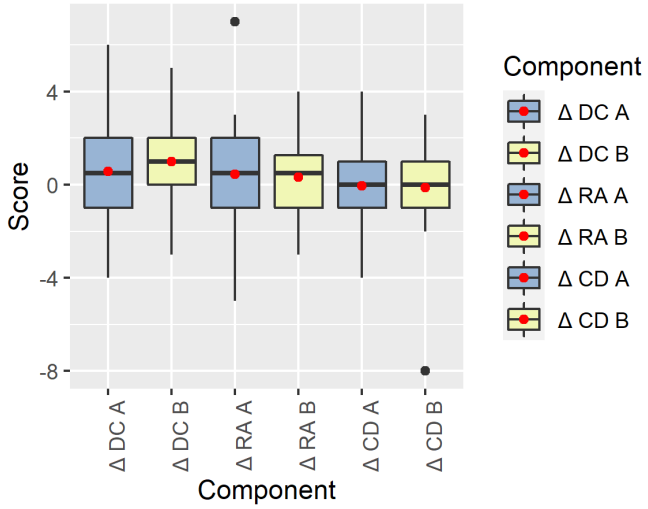


Figure 6.2: The distributions of the score differences of the M-GUDS-S in a box plot between version A (blue) and version B (yellow). The three components (DC, RA and CD) are on the x-axis. The difference score of each component is on the y-axis.

As mentioned before, the alternative hypothesis $H1.2_A$ reads: *The differences between the scores on all three components of the M-GUDS-S are higher in version A than in version B.*

The null hypothesis $H1.2_0$ reads: *The differences between the scores on all three components of the M-GUDS-S between version A and version B are equal.*

Therefore, the Wilcoxon rank-sum test is used to test H1.2 for each component, as the samples are independent. The results are as follows:

- The experimental DC differences were lower (median = 0.5, IQR = 3.0) than the control DC differences (median = 1.0, IQR = 2.0). According to a Wilcoxon rank-sum test (one-tailed) for independent samples, this difference was *not* significant, $W = 114$, $p = .30$.
- The experimental RA differences were higher (median = 0.5, IQR = 3.0) than the control RA differences (median = 0.5, IQR = 2.3). According to a Wilcoxon rank-sum test (one-tailed) for independent samples, this difference was *not* significant, $W = 129.5$, $p = .48$.
- The experimental CD differences were lower (median = 0.0, IQR = 2.0) than the control CD differences (median = 0.0, IQR = 2.0). According to a Wilcoxon rank-sum test (one-tailed) for independent samples, this difference was *not* significant, $W = 121.5$, $p = .60$.

Thus, $H1.2A_A$ is rejected for all three components, as the p-values were all not significant. Moreover, the differences in DC even turned out to be in the opposite direction of what was hypothesised.

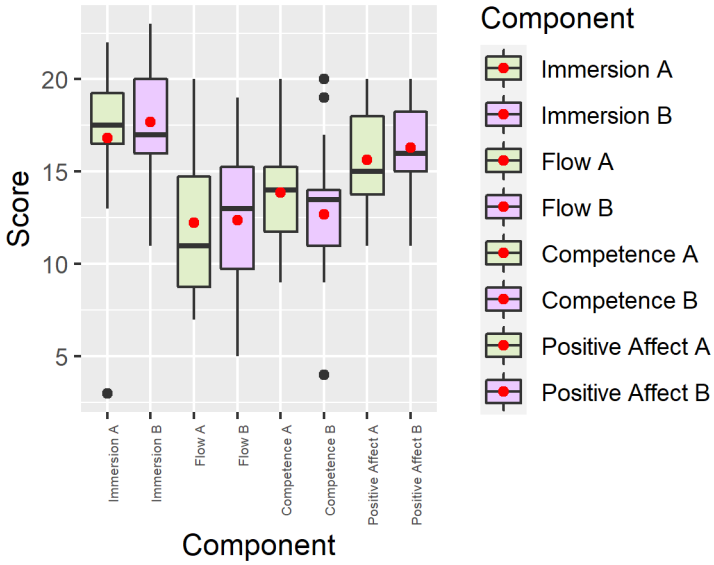


Figure 6.3: The distributions of the GEQ in a box plot between version A (green) and version B (purple). The four components (Immersion, Flow, Competence and Positive Affect) are on the x-axis. The score is on the y-axis.

6.1.1.3 Summary of the M-GUDS-S data

Looking at the whole sample, only post-test DC scored significantly higher than pre-test DC. As hypothesised, post-test RA scored higher as well, while post-test CD yielded lower scores. Nevertheless, the results from the other two components were not significant. When splitting the data into two groups, version A (the experimental group) and version B (the control group), the score differences (Δ) were very small, hence these tests yielded no significant results either. Nonetheless, the Δ DC_A differences were even lower then the Δ DC_B, which is the opposite of what was expected.

6.1.1.4 Differences between versions in GEQ

Figure 6.3 displays the distributions of the four components of the GEQ in a box plot, where each component is compared between version A (green) and version B (purple). This is visible on the x-axis. The score for each component is on the y-axis. Additionally, the differences in score between both versions are small. However, in some box plots the mean lies a little bit further away from the median. This indicates a skew in the data. Moreover, for both versions, Immersion scores the highest. This is followed by Positive Affect and then Competence. Flow scores the lowest out of the four components.

To reiterate, the alternative hypothesis $H2_A$ is presented as: *Version A of the game scores higher on all four components of GEQ than version B.*

The null hypothesis $H2_0$ is presented as: *There are no differences between the scores on all four components of the GEQ between version A and version B.*

Hence, the Wilcoxon rank-sum test is used to test H2 for each component, as the samples are independent. The results are as follows:

- The scores for Immersion in the experimental condition were higher (median = 17.5, IQR = 2.8) than the control condition (median = 17.0, IQR = 4.0). According to a Wilcoxon rank-sum test (one-tailed) for independent samples, this difference was *not* significant, $W = 124.5$, $p = .55$.
- The scores for Flow in the experimental condition were lower (median = 11.0, IQR = 6.0) than the control condition (median = 13.0, IQR = 5.5). According to a Wilcoxon rank-sum test (one-tailed) for independent samples, this difference was *not* significant, $W = 123.5$, $p = .43$.
- The scores for Competence in the experimental condition were higher (median = 14.0, IQR = 3.5) than the control condition (median = 13.5, IQR = 3.0). According to a Wilcoxon rank-sum test (one-tailed) for independent samples, this difference was *not* significant, $W = 147$, $p = .23$.
- The scores for Positive Affect in the experimental condition were lower (median = 15.0, IQR = 4.3) than the control condition (median = 16.0, IQR = 3.3). According to a Wilcoxon rank-sum test (one-tailed) for independent samples, this difference was *not* significant, $W = 108$, $p = .22$.

Thus, $H2_A$ is rejected for all four components, as all p-values turned out to be not significant. Furthermore, the scores for Flow and Positive Affect were even oriented in the opposite direction as expected.

6.1.1.5 Identification with a marginalised group in GEQ

The demographic questionnaire asked participants if they identified with a marginalised group, as it is hypothesised that this control variable could positively influence the gaming experience. As stated before, 10 participants answered "Yes", 19 participants replied "No" and 3 participants said "I don't know/Prefer not to say". Since the latter group fits neither the binary category of "Yes/No", we decided to discard these 3 responses in the analysis, as we are focusing on the effect of marginalisation. Hence, the data is split into two parts. This means that 10 participants belong to the marginalised group and 19 participants do not. Note that the sample sizes are unequal.

Figure 6.4 displays the distributions of the four components of the GEQ in a box plot, where each component is compared between the Y group (green) and the N group (purple). This can be seen on the x-axis, while the score is visible on the y-axis. Furthermore, the differences between both groups are visible. The N group scores higher on every component, which is completely the opposite of what was hypothesised. Similarly to the previous section, for both groups, Immersion scores the highest, which is followed by Positive Affect and then Competence. Flow scores the lowest out of the four components.

To repeat, the alternative hypothesis $H3_A$ represents the following: *Players who identify with a marginalised group score higher on all four components of GEQ compared to players who do not.*

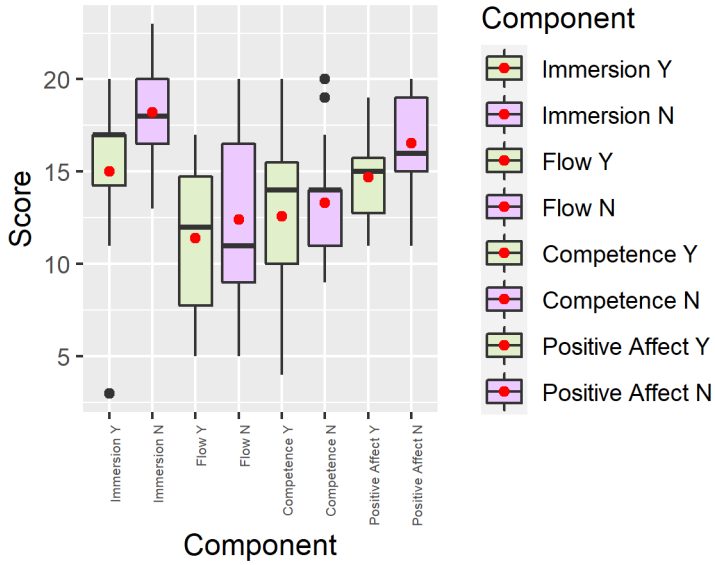


Figure 6.4: the distributions of the GEQ in a box plot per group, so identification with a marginalised group (Y, green) and no identification with a marginalised group (N, purple). The four components (Immersion, Flow, Competence and Positive Affect) are on the x-axis. The score is on the y-axis.

The null hypothesis $H3_0$ represents the following: *There are no differences between the scores on all four components of the GEQ between players who identify with a marginalised group and players who do not.*

Thus, the Wilcoxon rank-sum test is used to test H2 for each component, as the samples are independent. The results are as follows:

- The scores for Immersion in the marginalised group were lower (median = 17.0, IQR = 2.8) than the non-marginalised group (median = 18.0, IQR = 3.5). According to a Wilcoxon rank-sum test (one-tailed) for independent samples, this difference was significant, $W = 55.5$, $p = .03$. Furthermore, it represented a strong effect, $r = -.67$.
- The scores for Flow in the marginalised group were lower (median = 12.0, IQR = 7.0) than the non-marginalised group (median = 11.0, IQR = 7.5). According to a Wilcoxon rank-sum test (one-tailed) for independent samples, this difference was *not* significant, $W = 83$, $p = .29$.
- The scores for Competence in the marginalised group were higher (median = 14.0, IQR = 5.5) than the non-marginalised group (median = 14.0, IQR = 3.0). According to a Wilcoxon rank-sum test (one-tailed) for independent samples, this difference was *not* significant, $W = 97$, $p = .46$.
- The scores for Positive Affect in the marginalised group were lower (median = 15.0, IQR = 3.0) than the non-marginalised group (median = 15.0, IQR = 4.0). According to a Wilcoxon rank-sum test (one-tailed) for independent samples, this difference was significant, $W = 58.5$, $p < .05$. Furthermore, it represented a strong effect, $r = -.63$.

The p-values for Flow and Competence turned out to be not significant. Additionally, the scores for Immersion, Flow and Positive Affect were even oriented in the opposite direction, where both Immersion and Positive Affected contained significant p-values. However, the significance of these p-values do not support $H3_A$, but rather provide evidence to reject it, as they support lower scores instead of higher ones. Nevertheless, this also implies that they provide evidence to reject $H3_0$ as well, since they imply a difference exists. Therefore, $H3_A$ is rejected for all four components.

6.1.1.6 Summary of the GEQ data

Looking at version A (the experimental group) versus version B (the control group), the groups did not score significantly different on all four components. However, the scores for Flow and Positive Affect were both lower in version A. This was the inverse of what was predicted. Regarding the Y group (marginalised) vs N group (not marginalised), there were two significant results. First, the Y group scored significantly lower on Immersion than the N group. This was not as expected. Second, the Y group scored significantly lower on Positive Affect. This was also not predicted. Additionally, the Y group scored lower on Flow as well. Again, the opposite was hypothesised, but the latter difference was not significant. Lastly, the Y group scored higher on Competence. Yet, this was not significant.

6.1.2 Exploratory research

This part of the analysis falls under exploratory research, as no hypotheses were made a priori. Yet, statistical tests were applied to check for bias in the data. Moreover, the descriptive statistics and the box plots are shown and elaborated upon. Again, all box plots indicated some outliers. Nevertheless, these were not removed, as they could represent some novelties instead of deviations.

6.1.2.1 Identification with a marginalised group in M-GUDS-S

Continuing with the same two groups as we have analysed in H3, we decided to analyse if identification with a marginalised group could have any effect on the differences in the pre-test and post-test M-GUDS-S scores.

Figure 6.5 shows the distributions of the differences of each of the three components of the M-GUDS-S in a box plot, where Y (blue) is compared to N (yellow). This can be seen on the x-axis. On the y-axis, the score difference is represented. Although results across the groups do not differ much from each other, the bigger IQR and whiskers for the Y group indicate that their answers varied more than in the N group. Moreover, the mean and the median in each box plot lie close to zero. This suggests that the participants' answers to the M-GUDS-S did not vary much after playing the game, in general. Additionally, in both groups, DC and RA have a slight positive difference, while the CD difference is closer to zero.

To check for bias, the Wilcoxon rank-sum test is used, as the samples are independent. The results are as follows:

- The scores for Δ DC in the marginalised group were higher (median = 0.5, IQR = 2.5) than the non-marginalised group (median = 1.0, IQR = 3.0). According to a Wilcoxon rank-sum test (two-tailed) for independent samples, this difference was *not* significant, $W = 107.5$, $p = .56$.

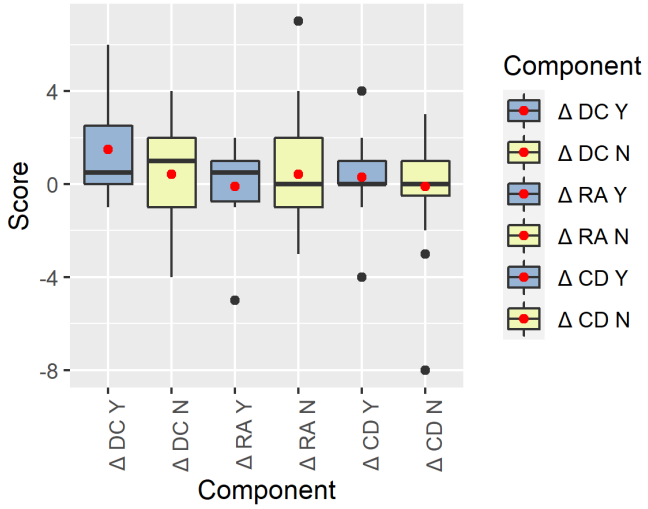


Figure 6.5: The distributions of the score differences of the M-GUDS-S in a box plot per group, so identification with a marginalised group (Y, blue) and no identification with a marginalised group (N, yellow).The three components (DC, RA and CD) are on the x-axis. The difference score of each component is on the y-axis.

- The scores for Δ RA in the marginalised group were lower (median = 0.5, IQR = 1.8) than the non-marginalised group (median = 0.0, IQR = 3.0). According to a Wilcoxon rank-sum test (two-tailed) for independent samples, this difference was *not* significant, $W = 90.5$, $p = .83$.
- The scores for Δ CD in the marginalised group were higher (median = 0.0, IQR = 1.0) than the non-marginalised group (median = 0.0, IQR = 1.5). According to a Wilcoxon rank-sum test (two-tailed) for independent samples, this difference was *not* significant, $W = 98$, $p = .89$.

Hence, all p-values are not significant, which implies that there is no bias in the data.

6.1.2.2 Controlling for Gender

Focusing on another control variable, the demographic questionnaire asked participants for their gender. As mentioned before, 24 participants answered "Female" and 8 participants replied "Male". Thus, the data is divided into two parts. This means that 24 participants belong to the female group (F) and 8 participants fit in the male group (M). Note that the sample sizes are unequal.

First, we analysed the influence of gender on the differences in the pre-test and post-test M-GUDS-S. Figure 6.6 presents the distributions of the differences of each of the three components of the M-GUDS-S in a box plot, where Females (blue) are compared to Males (yellow). This can be seen on the x-axis. On the y-axis, the score difference is represented. Here, the results are visibly different between the groups. For instance, DC_M scores higher than DC_F. The box plot for CD_M looks surprising, as its shape could indicate that the participants belonging to this group had very similar answers. This is less the case for CD_F, where the shape of the box plot indicates the answers varied more. In general, the mean and

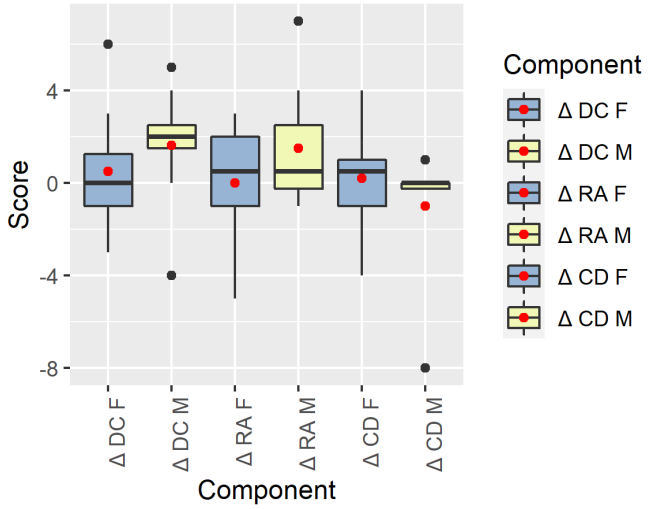


Figure 6.6: The distributions of the score differences of the M-GUDS-S in a box plot between Females (blue) and Males (yellow). The three components (DC, RA and CD) are on the x-axis. The difference score of each component is on the y-axis.

the median in each box plot lie close to zero. This suggests that the participants' answers to the M-GUDS-S did not vary much after playing the game. Additionally, in both groups, DC and RA have a slight positive difference, while the CD difference is closer to being below zero.

To test for bias, the Wilcoxon rank-sum test is used, as the samples are independent. The results are as follows:

- The scores for Δ DC in the female group were lower (median = 0.0, IQR = 2.3) than the male group (median = 2.0, IQR = 1.0). According to a Wilcoxon rank-sum test (two-tailed) for independent samples, this difference was *not* significant, $W = 58$, $p = .09$.
- The scores for Δ RA in the female group were lower (median = 0.5, IQR = 3.0) than the male group (median = 0.5, IQR = 2.8). According to a Wilcoxon rank-sum test (two-tailed) for independent samples, this difference was *not* significant, $W = 71$, $p = .27$.
- The scores for Δ CD in the female group were higher (median = 0.5, IQR = 2.0) than the male group (median = 0.3, IQR = 1.5). According to a Wilcoxon rank-sum test (two-tailed) for independent samples, this difference was *not* significant, $W = 120$, $p = .28$.

Since all p-values are not significant, it implies that there is no bias in the data.

Next, the influence of gender on the GEQ is analysed. Figure 6.7 displays the distributions of the four components of the GEQ in a box plot, where each component is compared between Females (green) and Males (purple). The differences between both groups are visible. For

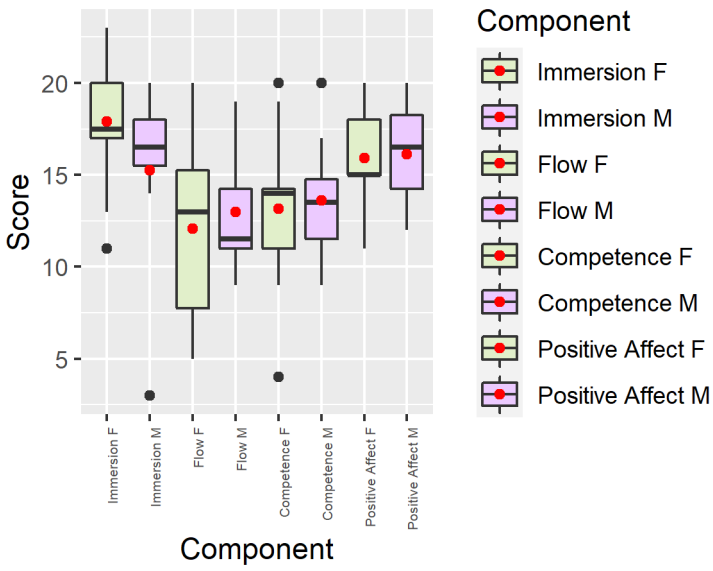


Figure 6.7: The distributions of the GEQ in a box plot between Females (green) and Males (purple). The four components (Immersion, Flow, Competence and Positive Affect) are on the x-axis. The score is on the y-axis.

instance, the IQR range for Flow in the Female group is larger than in the Male group. This could imply that there is more variety in the scores in the first group. Likewise, the Immersion in the Female group seemed to be higher than in the Male group. While Competence has very comparable scores in both groups, Positive Affect tends to score slightly higher in the Male group. Similarly to the previous sections, Immersion scores the highest, followed by Positive Affect and then Competence. Flow scores the lowest out of the four components.

To test for bias, the Wilcoxon rank-sum test is used, as the samples are independent. The results are as follows:

- The scores for Immersion in the female group were higher (median = 17.5, IQR = 3.0) than the male group (median = 16.5, IQR = 2.0). According to a Wilcoxon rank-sum test (two-tailed) for independent samples, this difference was *not* significant, $W = 127.5$, $p = .17$.
- The scores for Flow in the female group were higher (median = 13.0, IQR = 7.5) than the male group (median = 11.5, IQR = 3.3). According to a Wilcoxon rank-sum test (two-tailed) for independent samples, this difference was *not* significant, $W = 85$, $p = .63$.
- The scores for Competence in the female group were higher (median = 14.0, IQR = 3.3) than the male group (median = 13.5, IQR = 3.3). According to a Wilcoxon rank-sum test (two-tailed) for independent samples, this difference was *not* significant, $W = 97$, $p = .96$.
- The scores for Positive Affect in the female group were lower (median = 15.0, IQR = 3.0) than the male group (median = 16.5, IQR = 4.0). According to a Wilcoxon

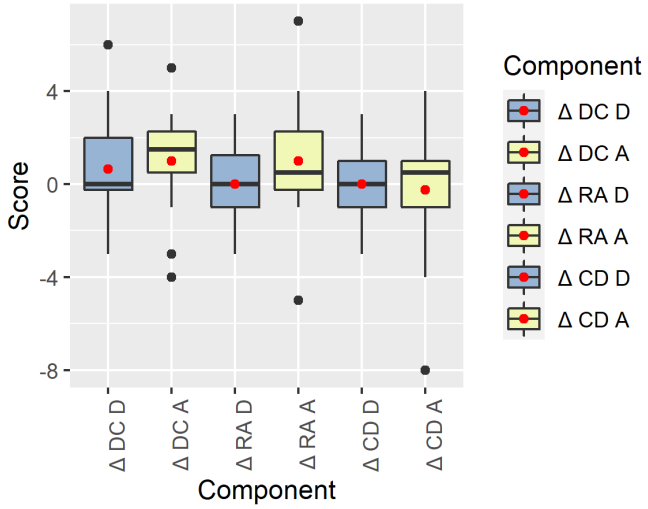


Figure 6.8: The distributions of the score differences of the M-GUDS-S in a box plot between Dutch (blue) and Abroad (yellow). The three components (DC, RA and CD) are on the x-axis. The difference score of each component is on the y-axis.

rank-sum test (two-tailed) for independent samples, this difference was *not* significant, $W = 88.5$, $p = .74$.

Due to the fact that all p-values are not significant, it indicates that there is no bias in the data.

6.1.2.3 Controlling for Nationality

Focusing on another control variable, the demographic questionnaire asked participants for their nationality. To reiterate, 20 participants answered "Dutch" and 12 participants filled in another country. To make a binary diversion, all non-Dutch participants are categorised under the "Abroad" group. Thus, the data is divided into two parts. This means that 20 participants belong to the Dutch group (D) and 12 participants fit in the Abroad group (A). Note that the sample sizes are unequal.

First, we analysed the influence of nationality on the differences in the pre-test and post-test M-GUDS-S. Figure 6.8 shows the distributions of the differences of each of the three components of the M-GUDS-S in a box plot, where Dutch (blue) is compared to Abroad (yellow). This can be seen on the x-axis. On the y-axis, the score difference is represented. The results differ somewhat between the groups. For example, the Abroad group tends to score a little bit higher on all three components. In general, the mean and the median in each box plot lie close to zero. This suggests that the participants' answers to the M-GUDS-S did not vary much after playing the game. Additionally, in both groups, DC and RA have a slight positive difference, while the CD difference is closer to being below zero.

To test for bias, the Wilcoxon rank-sum test is used, as the samples are independent. The results are as follows:

- The scores for Δ DC in the Dutch group were lower (median = 0.0, IQR = 2.3) than the Abroad group (median = 1.5, IQR = 1.8). According to a Wilcoxon rank-sum test

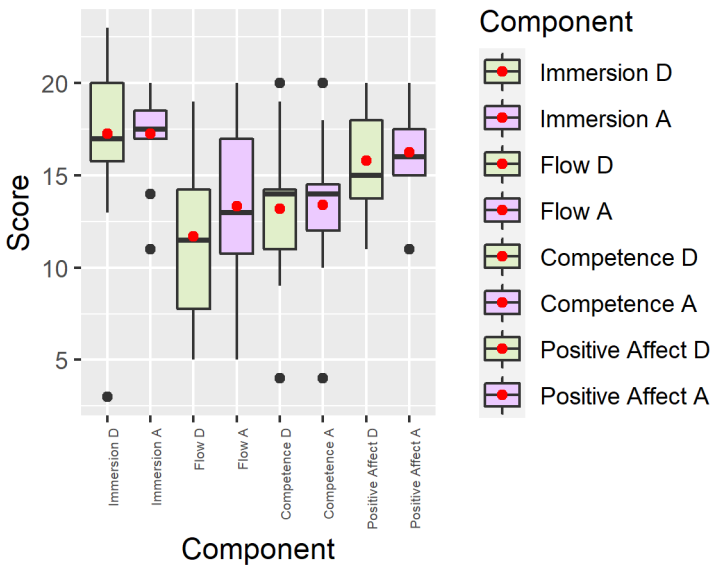


Figure 6.9: The distributions of the GEQ in a box plot between Dutch (green) and Abroad (purple). The four components (Immersion, Flow, Competence and Positive Affect) are on the x-axis. The score is on the y-axis.

(two-tailed) for independent samples, this difference was *not* significant, $W = 94.5$, $p = .31$.

- The scores for Δ RA in the Dutch group were lower (median = 0.0, IQR = 2.3) than the Abroad group (median = 0.5, IQR = 2.5). According to a Wilcoxon rank-sum test (two-tailed) for independent samples, this difference was *not* significant, $W = 90$, $p = .24$.
- The scores for Δ CD in the Dutch group were lower (median = 0.0, IQR = 2.0) than the Abroad group (median = 0.5, IQR = 2.0). According to a Wilcoxon rank-sum test (two-tailed) for independent samples, this difference was *not* significant, $W = 110.5$, $p = .70$.

Because all p-values are not significant, it is suggested that there is no bias in the data.

Next, the influence of nationality on the GEQ is analysed. Figure 6.9 displays the distributions of the four components of the GEQ in a box plot, where each component is compared between Dutch (green) and Abroad (purple). Nevertheless, the differences between both groups are visible. For instance, the IQR range for both Immersion and Positive Affect in the Abroad group is smaller than in the Dutch group. This could indicate that there is more variety in the scores in the latter group. Likewise, the Flow in the Abroad group seemed to be higher than in the Dutch group. Lastly, Competence has very comparable scores in both groups. Similarly to the previous sections, Immersion scores the highest, which is followed by Positive Affect and then Competence. Flow scores the lowest out of the four components.

To check for bias, the Wilcoxon rank-sum test is used, as the samples are independent. The results are as follows:

- The scores for Immersion in the Dutch group were lower (median = 17.0, IQR = 4.3) than the Abroad group (median = 17.5, IQR = 1.5). According to a Wilcoxon rank-sum test (two-tailed) for independent samples, this difference was *not* significant, $W = 121$, $p = .97$.
- The scores for Flow in the Dutch group were lower (median = 11.5, IQR = 6.5) than the Abroad group (median = 13.0, IQR = 6.3). According to a Wilcoxon rank-sum test (two-tailed) for independent samples, this difference was *not* significant, $W = 95$, $p = .33$.
- The scores for Competence in the Dutch group were lower (median = 14.0, IQR = 3.3) than the Abroad group (median = 13.5, IQR = 2.5). According to a Wilcoxon rank-sum test (two-tailed) for independent samples, this difference was *not* significant, $W = 113$, $p = .78$.
- The scores for Positive Affect in the Dutch group were lower (median = 15.0, IQR = 4.3) than the Abroad group (median = 16.0, IQR = 2.5). According to a Wilcoxon rank-sum test (two-tailed) for independent samples, this difference was *not* significant, $W = 106.5$, $p = .59$.

Because all p-values are not significant, it implies that there is no bias in the data.

6.1.2.4 Summary of the exploratory M-GUDS-S data

There were no significant results obtained in the exploratory M-GUDS-S data. Regarding the Y group (marginalised) versus N group (not marginalised), the Y group scored higher for Δ DC and Δ CD, but lower for Δ RA. Looking at the Females versus the Males, the females scored lower for Δ DC and Δ RA, but higher for Δ CD. As for the Dutch versus Abroad, the Dutch scored lower on all three components.

6.1.2.5 Summary of the exploratory GEQ data

There were no significant results found in the exploratory GEQ data. Looking at the Females versus the Males, the females scored higher for Immersion, Flow and Competence, but they scored lower on Positive Affect. As for the Dutch versus Abroad, the Dutch scored lower on all four components.

6.1.3 Choice distribution flow

Table 6.2 displays the player statistics regarding the achieved endings. As can be seen, the majority of the players received the 'good' ending. Only one player reached the 'bad' ending. They explained that they made 'bad' choices on purpose, as they were curious to see what would happen in the story. They stated they would not show such behaviour in real life.

To illustrate the relationships between the dilemmas, as well as the choice distribution in each dilemma, a Sankey flow chart is made. This can be seen in Figure 6.10. The full size of the figure can be found in Appendix A.2, which is made in SankeyMATIC. The hashtag (#) denotes the dilemma, followed by a one-word label of it. The labels without a hashtag represent the choice options. The number after each label represents the number of players that encountered it. The chart shows absolute numbers, where 32 is the maximum number,

Table 6.2: The player statistics regarding the achieved endings.

Ending	**Number of participants**	**Percentage**
Good	20	63%
Neutral	11	34%
Bad	1	3%

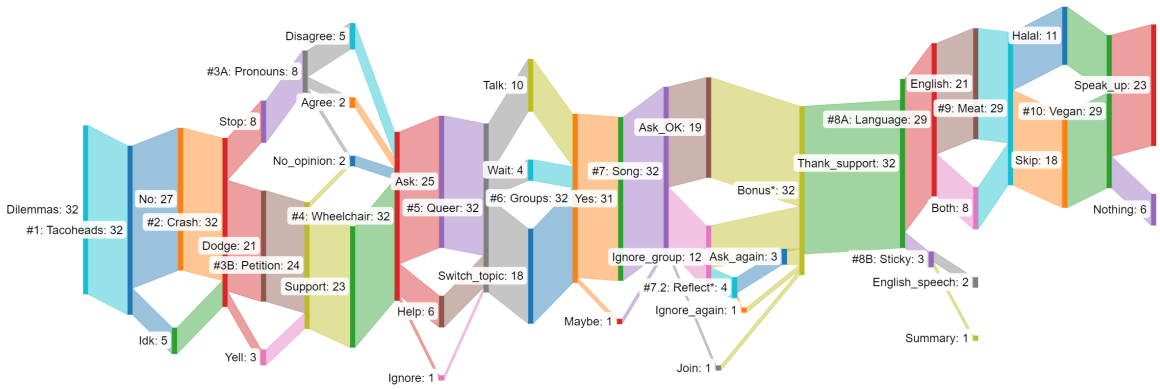


Figure 6.10: The Sankey flow chart of the choices players made and the dilemmas, with a hashtag (#), they encountered. The number after each label represents the number of players that encountered, where 32 is the maximum number. The width of a branch indicates the number of people that made a certain choice.

as this indicates the total number of participants. Note that there are two dilemmas with an asterisk (*). That is, these dilemmas had to be 'unlocked'. The first one is Dilemma #7.2, as it was only present in version A after choosing the 'Ignore_group' option in Dilemma #7. The second is the Bonus dilemma, present in both versions. This dilemma was only shown to players who had answered the bonus question correctly.

The Sankey flow chart shows that most players picked the same choices and that only a few players digressed, as the width of the branch indicates the amount. While the branches split after some choices and lead to a different dilemma, they reunited again with a common dilemma. This is done for the purpose of the research, as it was important to give the players a similar narrative experience, both in content and length. Yet, on the right part of the figure, it is visible that after Dilemma #8A, the player receives two extra dilemmas. After Dilemma #8B, the branch ends earlier. Only 3 out of the 32 players did not encounter these two extra dilemmas.

Additionally, Table 6.3 shows the Sankey flow chart in percentages instead of absolute numbers. The options are ordered 'Good-Neutral-Bad'. The options are not equally distributed, meaning that players seem to have a strong preference for certain choices. Usually these choices were the 'best' option, which implies that players felt what they 'should' do in each situation. This finding is further elaborated in section 6.3.2.1. Moreover, note that not each dilemma adds up to 100%, meaning that not all players encountered the same dilemmas. This

was dependent on the story branch and game version, as explained before. While most players either choose the 'good' or the 'neutral' option, this does not seem to hold for Dilemma #5. Here, the majority of the players chose the 'bad' option. We suspect that this is not due to malicious intent of the players, but rather a mismatch between expectations and outcomes, as the players did not perceive that option necessarily as bad.

Table 6.3: The percentage distributions of choices made in each dilemma. The choice options are ordered 'Good-Neutral-Bad'.

Dilemmas	**Choice options**			**Total**
1: Tacoheads	No (84%)	Idk (16%)	-	100%
2: Crash	Stop (25%)	Dodge (66%)	Yell (9%)	100%
3A: Pronouns	Disagree (16%)	No_opinion (3%)	Agree (6%)	25%
3B: Petition	Support (72%)	No_opnion (3%)	Sceptic (0%)	75%
4: Wheelchair	Ask (78%)	Help (19%)	Ignore (3%)	100%
5: Queer	Talk (31%)	Wait (13%)	Switch_topic (56%)	100%
6: Groups	Yes (97%)	Maybe (3%)	No (0%)	100%
7: Song	Ask (59%)	Ignore (38%)	Join (3%)	100%
7.2*: Reflect	Ask (9%)	Ignore (3%)	-	12%
Bonus*	Thank_support (88%)	Downplay (0%)	-	88%
8A: Language	English (66%)	Both (25%)	Dutch (0%)	91%
8B: Sticky	English (6%)	Summary (3%)	Dutch (0%)	9%
9: Meat	Halal (35%)	Skip (56%)	Lie (0%)	91%
10: Vegan	Disagree (72%)	Nothing (19%)	Agree (0%)	91%

6.2 Main findings of the quantitative analysis

As there was not enough evidence to confirm our hypotheses, all of them were rejected. Nevertheless, there were still three signficiant results. These are:

- Post-test DC scored *higher* than pre-test DC.
- The scores for Immersion in the marginalised group were *lower* than the non-marginalised group.
- The scores for Positive Affect in the marginalised group were *lower* than the non-marginalised group.

Moreover, the Sankey flow chart illustrates that the majority of the players chose the same choices and hence achieved the same ending.

6.3 Qualitative analysis

32 interviews were transcribed, with a total duration of 7 hours, 14 minutes and 23 seconds. Due to upload limitations, several tools have been used for transcription. The first 3 transcripts were manually transcribed with the help of oTranscribe[1], the last 5 transcripts were done in Otter.ai[2] and everything else was performed in Microsoft Word's automatic transcription tool. Then, all transcriptions were coded in NVivo version 14. A combination of a priori and emergent coding was used, however the focus was on emergent coding. The emergent coding used the Straussian Grounded Theory, which consists of three parts: 1) open coding, 2) axial coding, and 3) selective coding. This method allows for a systematic and interactive approach for analysis. Furthermore, the codes could support the theories that result from the data (Blandford et al., 2016; Corbin & Strauss, 1990).

6.3.1 Open and axial coding

During open coding, initial codes were generated from the transcriptions. A code could be created in two ways. First, a code was made beforehand, by adding it to the code list that was based on the interview questions. Then, interview snippets that fitted these codes were assigned to it. If a snippet did not fit into the a priori codes, the second method was applied. That is, new codes were created through emergent coding. In total, 104 codes were generated.

The next step consisted of axial coding, which is grouping the codes into hierarchical categories. The 104 codes that were identified in the previous step were categorised into bigger categories, until the main level was reached. As can be seen in Figure 6.11, these main categories act as parent codes and are as follows: 1) Narrative, 2) Dilemmas, 3) Visuals, 4) UX, and 5) Improvements. The circled number underneath each parent represents the number of child codes.

The full structure of each parent code is displayed in Figure 6.12. The complete hierarchical code structure made in Xmind is put in Miro and can be found here[3]. Similarly, the full size of Figure 6.12 can be found in Appendix A.3.

[1] https://otranscribe.com/
[2] https://otter.ai/
[3] https://miro.com/app/board/uXjVMDZebz8=/?share_link_id=701564651950

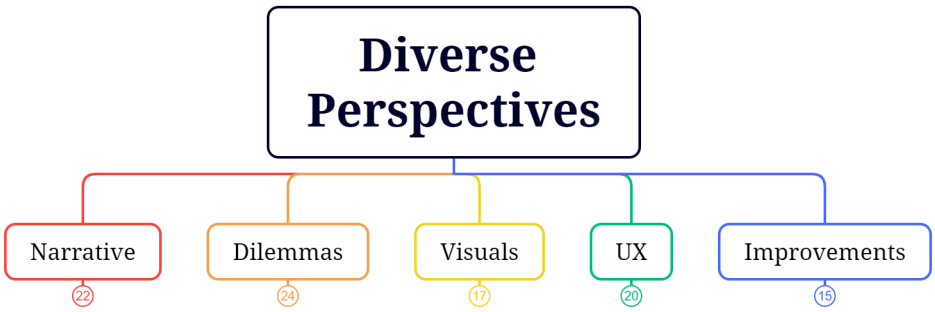


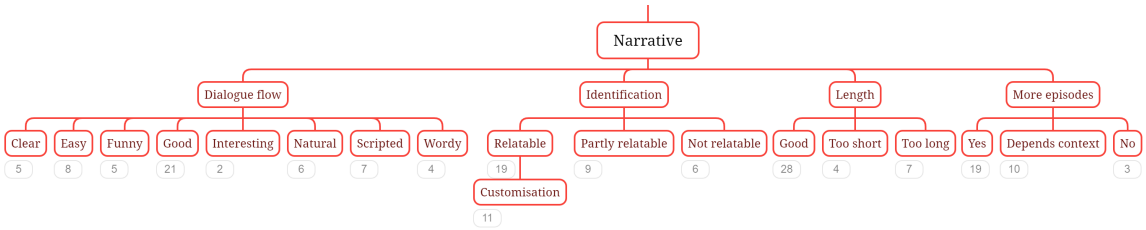
Figure 6.11: The main categories after open and axial coding. The circled number underneath each parent represents the number of child codes.

The number at the bottom of the child nodes indicates the number of participants that mentioned that specific code. This is to illustrate the prominence of each code. Note that the sum of the children can be less or more than 32 (which is the total number of participants). This is because some snippets were applicable to multiple codes, while others did not contain a specific code at all.

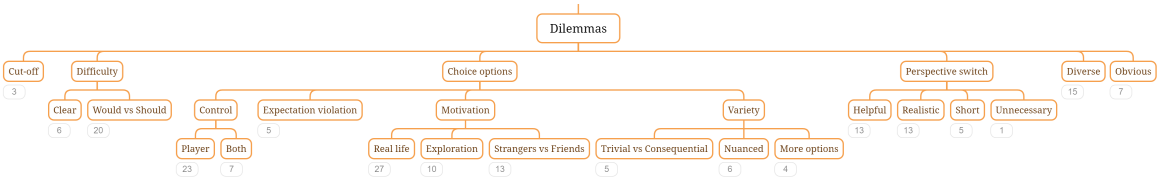
To illustrate how a priori coding was used in combination with emergent coding, the *Narrative* category is taken as an example. Each of the four subcategories of the *Narrative* parent code represents a question in the interview. Respectively, these questions were: "*What did you think of the dialogue flow?*", "*How much did you feel like you could identify with the characters?*", "*What did you think of the length of the story?*" and "*What do you think of the idea of more episodes?*". Thus, these subcategories were coded a priori. The reason why these questions were chosen to act as a priori codes in this specific category, is that they were all associated to the narrative. The same reason holds for the grouping of the other main categories. The answers to these questions became the child codes and the relationship between parents and children were formed during emergent coding.

The *Dilemmas* and *Visuals* main categories were created in a similar vein. For *Dilemmas*, we used the following questions for the a priori codes: "*What was your first impression of the game?*", "*What did you think of the dilemmas?*", "*What did you think of the perspective switch?*", "*Would you make the same choices in real life?*" and "*Did it feel like you were making your own choices or the avatar's choices?*". Regarding *Visuals*, the a priori codes were based on these questions: "*What did you think of the visuals?*" and "*How did you find the layout to navigate through the game?*".

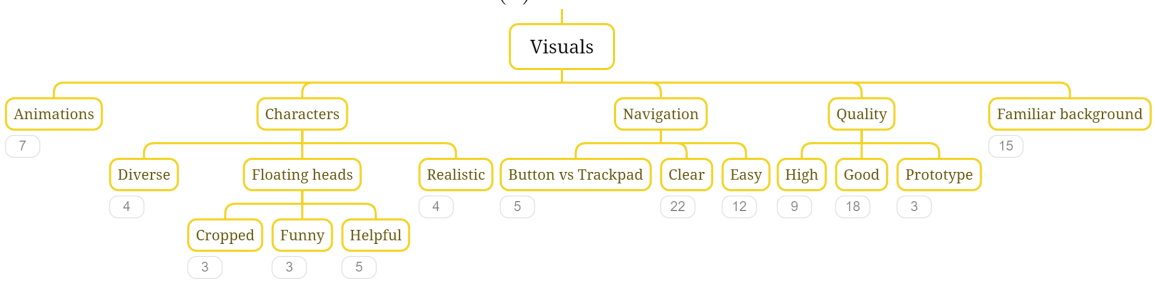
Emergent coding played a bigger role in the *UX* and *Improvements* main categories, as these questions could be answered very broadly and therefore only few codes could be made beforehand. Concerning *UX*, the codes that were made beforehand, were based on the following questions: "*What was your first impression of the game?*", "*Would it be useful in education about diversity?*" and "*Which ending did you get?*". For *Improvements*, the a priori code followed from this question: "*Do you have any other feedback?*".



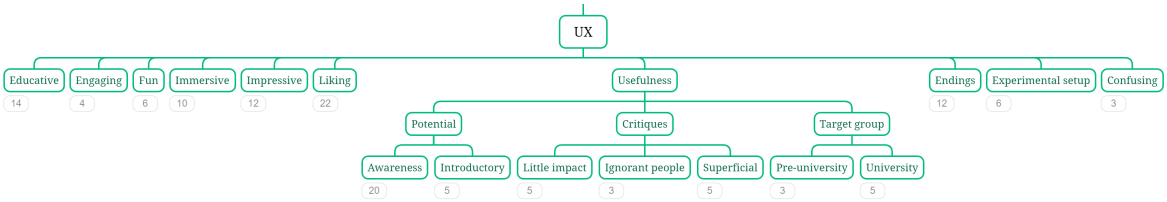
(a) Narrative



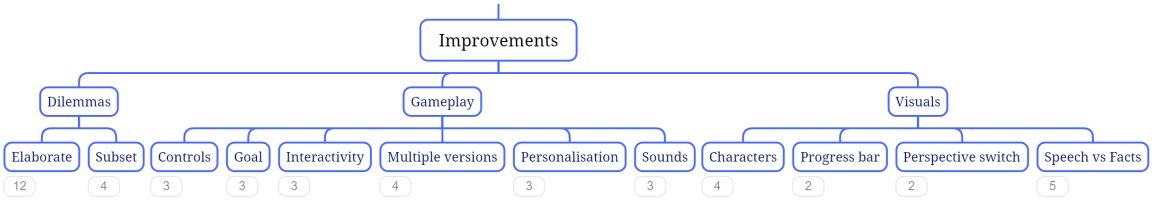
(b) Dilemmas



(c) Visuals



(d) UX



(e) Improvements

Figure 6.12: The five tree diagrams of the hierarchical code structure. The number at the bottom of the child nodes indicates the number of participants that mentioned that specific code.

Next, we elaborate on the hierarchical code tree that resulted from open and axial coding. Each main category is further depicted, by describing their child codes. Note that the *Improvements* category is explained in section 7.3.2.

6.3.1.1 Narrative

Narrative has four subcategories, namely: 1) Dialogue flow, 2) Identification, 3) Length, and 4) More episodes. Concerning *Dialogue flow*, participants mainly described it positively. For example, they found it 'easy', 'good' and 'natural'. A few participants found the dialogue 'scripted' and 'wordy'. Because the dialogue can make or break the narrative curve, the majority of positive opinions is seen as encouraging.

Regarding *Identification*, most participants found the story relatable. This ties back to the relatedness need of the SDT. Some participants did find it partly relatable, or not relatable at all. The participants that did find it relatable, remarked that the customisation of their avatar at the beginning of the game increased their relatability. Interestingly, a few participants customised their avatar with different features compared to their own situation in real life. For instance, they chose a different gender, nationality or both. They did this to see whether it would change their perspective on how they would perceive the situation. While one participant stated that it did change their decision-making, another participant reported that they did not experience any differences. In sum, this suggests customisation plays an important role in relatability. Relatability in turn can keep the player interested and engaged in the story, as it highlights the relevance of the material.

About *Length*, nearly all of the participants deemed it as appropriate. While a few participants found it too long, even less participants found it rather too short. This concludes us to perceive the length as good, it should not be any longer or shorter.

Relating to *More episodes*, the majority of the participants stated that they would be interested to play them if more were to be made. Only a few said they would not be interested and almost one third expressed that it would depend on the context of the story. If the story would have a strong narrative, they would probably be open to playing more episodes. This implies that *Diverse Perspectives* could have potential to expand.

6.3.1.2 Dilemmas

Dilemmas has a slightly more complicated structure than the previous main category. Therefore, the interpretations are more on a global level. Almost half of the participants found the dilemmas diverse, yet almost a quarter found them obvious. A few participants found some dilemmas cut-off, as they perceived them to end somewhat suddenly. This links to the expectation violation, where various participants mentioned that the characters would behave differently from what they expected. One participant remarked that this is an issue with visual novels in general, as two or three words in a choice option could never fully represent its consequences. Likewise, words could be interpreted in multiple ways that the researcher did not think of. In short, this implies that the dilemmas serve their function, however, there is still room for improvements.

Looking at the difficulty of the dilemmas, participants expressed that they doubted between behaving as they would do in real life or choosing the 'morally best' options. Yet, they were aware of what the game found morally best, might not always be realistic. This creates

a dilemma in itself. This an interesting development, as it really challenges the player to reflect on their decision-making.

When making a choice, most participants declared the control was in their hands. Meaning, it was fully their choice, instead of their avatar's. This links back to the *autonomy* component from the SDT. Several participants stated that it was a combination of both. That is, they took into account their avatar's perceived personality when making decisions. Despite the fact that most choices were made as participants would do in real life, some participants commented that they picked something just to explore what would happen next. Choosing a certain option also depended on whether strangers were involved in the situation or not. Moreover, some participants liked the nuanced level of the choices, while others would have liked more options. This shows that several factors influence decision-making. The choices offered were decent, yet there are still many opportunities to improve.

Lastly, the experimental group that did experience the perspective switch found it 'helpful' and 'realistic'. Some participants found it rather short, and only one participant did not find it necessary at all. This suggests that the switch achieved its intended effect of creating more understanding towards each other.

6.3.1.3 Visuals

Visuals has five subcategories, which are: 1) Animations, 2) Characters, 3) Navigation, 4) Quality, and 5) Familiar background. Regarding *Animations*, various participants mentioned that they liked the animations, such as the fading in and out of characters and backgrounds. They further remarked that it helped them with the flow of the story. Thus, the animations are considered as helpful.

Concerning *Characters*, a few participants stated that they were diverse and looked realistic. As described in section 4.2.4, characters were carefully selected with the intention of a diverse representation of realistic people. Some participants commented on the 'floating heads', as they called the pictures of the characters. Although a few participants found them awkwardly cropped off, others found them funny and helpful in highlighting who was talking at the moment. Hence, this implies that the visual representation of the characters served their intended function.

About *Navigation*, most participants found it 'clear' and 'easy'. Since *Diverse Perspectives* has a point-and-click format, the navigation is fairly simple and the layout looks similar to other visual novels. This was done to reduce the ECL as much as possible. Participants could use either the buttons or the trackpad. Both controls were explained before starting the game. While in a few cases participants accidentally activated some extra features that were not used in the game by misclicking or pressing the wrong buttons, this was quickly resolved. Therefore, navigating through the game is not perceived as a challenge.

Speaking of *Quality*, more than half of the participants found it 'good' and even a quarter found it 'high'. Three participants found it to be on 'prototype' level. While the quality of the visuals could be polished and increased, it is deemed more than sufficient for the current game.

Looking at *Familiar background*, a lot of participants mentioned that they recognised the backgrounds in the game and it helped them form a visual representation of the places where their character was. In other words, it increased their immersive experience. Thus, the backgrounds play a valuable role in the game.

6.3.1.4 UX

UX has a somewhat more complicated structure than the previous main category. Therefore, the interpretations are more on a global level. Generally speaking, participants had a positive UX while playing the game. This is suggested by words as 'engaging', 'fun' and 'immersive'. Moreover, many participants felt like they learned something from it, hence describing *Diverse Perspectives* as 'educative'. Only three participants reported that they were confused by the game at some point. Furthermore, quite a few participants were aware of the multiple endings, as they asked in the interview what the other endings were and how those narratives looked like.

Additionally, various participants said they liked the setup of the whole experiment, as they felt the questionnaires, the game and the interview complemented each other well. Especially the interview was seen as an opportunity to share their experiences and thoughts about the game. Hence, we are content with the conduction of the experiment.

While many participants saw *Diverse Perspectives* having potential in raising awareness and introducing people to diversity topics, some participants were also critical about the actual impact of the game and its narrative depth. Likewise, some participants doubted whether university students were the appropriate target group, as the game might be more useful for a younger or more conservative-minded audience. Although *Diverse Perspectives* is regarded as a useful and positive experience, further research highly likely benefits from taking the critiques into account as well.

6.3.2 Selective coding

The last step contained selective coding. It was analysed how the categories could be linked together and if any preliminary theories could be constructed. In total, eight preliminary theories were developed, including participant quotes to support them.

6.3.2.1 Would vs Should

The first theory was as follows: *When making a consequential choice, players seem to make a distinction between what they should do versus what they would actually do.* This approach created a dilemma in itself, as players stated that they had the sense there was a 'best' option, but that this option did not necessarily reflect their own actions. As P5, P8 and P11 explain:

> **P5:** "*Trying to figure out whether it should be what I would do or what I should do was something that I wondered about.*"
>
> **P8:** "*Maybe like for a lot of people, it's harder to answer the question because I think there's a difference in something that you want to do and the thing you really do if you were in a situation.*"
>
> **P11:** "*Yeah, that's why a lot of things that I selected was like doing nothing, because I usually don't do anything with that. Which is like, I know it, but that's what I do.*"

While most players chose options they would do in real life, some players liked to explore and chose the more 'provoking' options. This is mentioned by P4 and P26:

> **P4:** "*I do have to say I chose some things because I knew it's provocative. I just wanted to see what happens in the game. Yeah, I think that is the thing with games. I will choose something that I wouldn't choose in real life just because I just want to see if something weird happens.*"

> **P26:** "*I actually really like that about serious games, because you can just do whatever you want and without any consequence, or well in-game consequence, but not in real life consequence.*"

Hence, although players can have various motives to pick a certain option, they seemed to think and reason about each decision. Reflection about the content is desired, as it aims to engage the player with the material at a deeper level.

6.3.2.2 Strangers vs Friends

Because the story puts the player in the role of a new student, almost all dilemmas involve strangers. However, the second theory is phrased as: *Players are more likely to confront and comment on the behaviour of friends than strangers.* This is explained by P15 and P31:

> **P15:** "*Simply because I'm also not really a confrontational person, I would say. I don't feel comfortable going up to a group of people that is fighting or something and just inserting myself into that.*"

> **P31:** "*If it comes up with a friend or someone, who I think "OK, I don't agree with them". But yeah, I know them, then I can say it. But yeah, just random people feels a bit weird.*"

Therefore, how players would interact with friends and family is interesting to investigate in further research, since most dilemmas also seem to require the player to have enough confidence to approach strangers, regardless of the topic.

6.3.2.3 Trivial choices are not trivial

The third theory states: *The trivial choices provide a nice break from the consequential choices by lowering the 'moral pace'. Moreover, they give the player the opportunity to shape the personality of their avatar, which in turn could increase the bond between the player and the avatar.* As P7 and P17 commented:

> **P7:** "*And you could make silly choices. Like what food you eat. So that was really good, I think so it feels more like you can shape the character.*"

> **P17:** "*I liked that there were dilemmas that didn't really matter [...] It just makes it a little bit easier. Yeah, but it's not every choice has to be this important.*"

Unlike their label, trivial choices do not play a trivial role. They seem to be at least as important as the consequential choices, as they provide variety to the player's actions.

6.3.2.4 Increase emotional impact

The fourth theory suggests: *Increasing the emotional connection between the player and the characters is likely to increase the emotional impact of the dilemmas.* As P4 and P19 explain:

> **P4:** "*It can draw in people very deeply into the story and then make them very emotionally attached to characters and then once they are attached to it, I think it has huge potential to make them see the problems that they suffer.*"

> **P19:** "*I think more episodes could be beneficial because then you could also maybe go into depth if you have a more real friendship, because these two people just met. And how do you deal with certain dilemmas if you have like a real connection?*"

While the majority of the players regard the current length of the game as good and some stated the length should not be any longer, the game is likely to have potential to expand if it contained new content. Furthermore, creating more episodes appears to be beneficial, as it allows a deeper dive into the narrative and character development, which could increase the bond between the player and their avatar. A stronger connection is likely to increase the emotional impact of the dilemmas. Currently, the dilemmas were regarded more as an introduction to certain topics.

6.3.2.5 Perspective switch stimulates reflection

The fifth theory implies: *The perspective switch provides the player with an extra layer of information and brings nuance and reflection to the situation.* As P5 and P29 said:

> **P5:** "*The switching perspectives was nice because I felt like you could sort of get that understanding of the other person's perspective and get sort of an immediate reflection on what the choice you made was and how it affected them.*"

> **P29:** "*It's really interesting to get to know better what they're thinking at that point. And even though, like with Chad, I wasn't agreeing on what he said and his thoughts, but it's interesting then what he thinks. Even though I strongly disagree, but it's just, you get to know more of their backgrounds and what they're thinking at that point.*"

Although the perspective switch is perceived as a bit short, it helps the player understand and reflect on choices, both from their side as well as the NPC they interact with. This feature was seen as adding an extra dimension to the story, as it would be represented from both sides.

6.3.2.6 Raising awareness

The sixth theory expresses: *Ignorant and extreme people are probably not affected by the game.* As P14 and P27 mention:

> **P14:** "*If someone is ignorant, I don't know if they would be like "Oh, let me play this game".*"

P27: "*With the game you're never gonna change the minds of very extreme people.*"

However, the target group could shift to first-year university students during introduction week, or even teenagers, as remarked by P2 and P6:

P2: "*I think it would be nice if you for example in the introduction week or something asked students to play this game. To give them also the chance to think about it.*"

P6: "*I would actually think something like this would be better implemented at an earlier stage as opposed to at a university level stage. To me, that would make more sense, especially because when you consider the degree of maturity that people have in, for example, high school, I could see something like this being implemented or even middle school.*"

This is because the game is seen as an engaging way to raise awareness and interact with the topics. This is illustrated by P1 and P18:

P1: "*I think it can make people aware of differences and maybe pay more attention of differences around them and why people do such things. And speaking up more often [...] Because you see that people actually like that.*"

P18: "*It's important for people to be put in uncomfortable situations and act well.*"

Therefore, a suited target group could be the so-called 'middle group'. This consists of people who do not have formed a strong opinion on these topics yet, but are open to learn more about them. As the game could provide an opportunity to bring people into contact with these topics for the first time in a safe environment, *Diverse Perspectives* could be more effective at a younger target group.

6.3.2.7 Relevance

The seventh theory states: *Making the narrative Utrecht-centred fits the learner's world, and the fun facts were regarded as educative for players who have been living in Utrecht for a longer time, which increases the relevance and immersion of the game.* This is supported by P5, P17 and P30:

P5: "*I feel like it's infused with this value of, like getting to know Utrecht.*"

P17: "*I learned some new things about Utrecht. Although I'm living here for three years.*"

P30: "*I like that you used pictures of Utrecht, so you can really live in the game. If you know Utrecht a little bit, you know where you are at that moment.*"

Since the target group were UU students, having the game revolve around the life of a UU student increased the relatability of the story. This seems to have worked out well, as even players who have been living in Utrecht for quite a while still learned something from the game. In other words, *Diverse Perspectives* was not only useful to players who were new in Utrecht. In a similar vein, adapting the narrative to other cities seems to be promising, as the game mechanics stay the same.

6.3.2.8 Replay

The eight theory declares: *Players seem to be motivated to replay the game to explore the other options or reach the 'best' ending in case they failed to do so in their first play.* As P3 and P26 commented:

> **P3:** "*I guess the fact that I didn't win, I didn't get the best outcome kind of leaves me wanting to see, like OK, maybe if I do it again, do a couple of things differently, maybe I can succeed.*"
>
> **P26:** "*If I would play this game again, I think it would be fun to just answer completely the opposite, or just a little bit differently, to see what could happen.*"

By replaying the game, players can discover other paths until they have reconstructed the full narrative and unlocked all content. Replays also allow players to make connections between actions and consequences on a bigger scale, as they know more about the game world. However, the other side of the coin is that once players have received the 'best' ending, it is possible that they are not motivated to replay it. This is because they do not think there is anything to gain from the game anymore.

6.4 Main findings of the qualitative analysis

After using a combination of a priori and emergent coding, five main main categories emerged. These are: 1) Narrative, 2) Dilemmas, 3) Visuals, 4) UX, and 5) Improvements. The last catogry is explained in section 7.3.2. Each main category has several (grand)child codes.

Regarding *Narrative*, most participants found it 'positive' and 'relatable'. Its length was perceived as appropriate, while the majority of the participants would be interested in playing more episodes.

Concerning *Dilemmas*, they serve their function. Yet, there are still opportunities for improvement. Choosing between what one *would* do versus what one *should* do represented the core of each dilemma. Furthermore, the involvement of strangers in a situation acts as an influential factor as well. Nevertheless, most participants stated that they were in control of their own choices and not the avatar. Almost all of the participants who experienced a perspective switch, deemed it of additional value.

Looking at *Visuals*, the animations were liked by the participants. Additionally, the cast of characters in the game was perceived as diverse and realistic, while the familiar backgrounds increased the immersive experience. Although the quality of the visuals could be more polished, it was seen as more than sufficient for the current game. Lastly, the navigation was experienced as 'clear' and 'easy'.

With reference to *UX*, most participants experienced it as positive. Various participants commented positively on the setup of the whole experiment as well. While *Diverse Perspectives* has potential in raising awareness of and familiarising people with diversity topics, its actual impact and narrative depth should be researched further in order to truly contribute to these topics.

After analysing how the categories could be possibly linked together, eight preliminary theories were formed. These are: 1) Would vs Should, 2) Strangers vs Friends, 3) Trivial choices are not trivial, 4) Increase emotional impact, 5) Perspective switch stimulates reflection, 6) Raising awareness, 7) Relevance, and 8) Replay.

Chapter 7

Discussion

The aim of this research was to encourage inclusive behaviour among UU students through the means of *Diverse Perspectives*. In the field of serious games, a gap was identified of the combination of diversity with (international) student inclusion. By following a user-centred and iterative approach, *Diverse Perspectives* was developed and then tested by the target users. This section describes the interpretations of the results. Moreover, it elaborates on the limitations of this research and presents various directions for further research.

7.1 Interpretation of the results

7.1.1 Quantitative results

7.1.1.1 Scores from the M-GUDS-S and GEQ

Although none of our hypotheses could be confirmed, the results indicate that further research seems to be fruitful. R1 aimed to investigate the impact of the perspective switch on one's attitude towards diversity. Quantitatively, the perspective switch did not cause any significant differences and therefore it is likely that its impact is small. This is elaborated by H1.1 and H1.2 below. More on the switch can be found in section 6.3.1.2.

Regarding H1.1, both post-test DC and post-test RA scored higher, while post-test CD scored lower. This was as expected and the score differences could be explained by the fact that people are not perfectly consistent when filling in the same questionnaire twice. Especially when there is a time span of a week between them. This seems plausible. However, it could also mean that *Diverse Perspectives* did have an influence on the post-test scores. Nevertheless, only post-test DC showed a significant difference, with a medium effect size, while post-test RA and post-test CD did not. Moreover, both DC and RA had very similar scores in both tests, in general. Therefore, it is unclear to which degree the game influenced the scores.

While H1.2 did not have any significant results, it is interesting to notice that the differences in DC were even lower, instead of higher. This could be due to the small sample size. Further research should be conducted to clarify the cause of this unexpected outcome. Moreover, the results suggest that the perspective switch did not have a significant influence on the change in one's attitude towards diversity.

R2 intended to explore which factors could contribute to a positive gaming experience. In total, four components from the GEQ questionnaire were identified, namely: 1) Immersion, 2) Flow, 3) Competence, and 4) Positive Affect. However, concerning H2, no significant results were obtained either. This again implies the differences were small in the quantified data about the gaming experience between both versions. Flow and Positive Affect even had opposite results as hypothesised. Yet, in general, Immersion scored the highest, followed by Positive Affect, then Competence and at last Flow. Besides the fact that Immersion had one item more and hence had a higher maximum score, this seems somewhat peculiar. This is because it is likely that if players felt highly immersed, they would be highly in their flow as well. However, it is unclear yet why this does not seem to be the case. Again, this invites further studies to investigate how this could occur.

A reason why the version with the perspective switch did not contribute significantly to the gaming experience, could be that the switches were rather short. Moreover, its frequency depended on the player's choices, with a minimum of one switch and a maximum of seven. In other words, it could have been that the differences between the two versions was not as big as estimated and therefore no significant results could be found.

R3 targeted to study the effect of identifying with a marginalised group on the gaming experience. Concretely, the same four components of the GEQ that were studied in R2 were used. Looking at H3, Immersion and Positive Affect had significant differences. However, these differences turned out to be in the contrary direction as hypothesised. In other words, they provided evidence to reject both $H3_A$ as well as $H3_0$. Furthermore, the effect sizes represented strong effects. This indicates that people who do *not* identify with a marginalised group, significantly felt more immersed in the game and experienced more positive feelings. This same group also scored higher on Flow, but this difference was not significant.

One explanation why people who do identify with a marginalised group seem to have a lesser gaming experience overall, may be because the topics feel quite personal and could invoke mixed feelings. This could result in lower scores for Positive Affect. It could also reduce their flow and the immersive effect, as people could start thinking about their own experiences. That is, their mind could be drifting away from the game. Yet, this group scored higher on Competence. This may be explained by the fact that the scenarios could be seen as familiar and hence people could have felt more skilled in handling them.

The exploratory research indicated that identification with a marginalised group did not seem to have any influence on the pre-test versus post-test M-GUDS-S scores. A reason for this might be that identifying with this group does not necessarily mean that one's diversity attitude is more open-minded by default. Besides that, gender and nationality do not seem to influence the scores either. Future works should clarify whether this holds in general.

While no significant results were found for gender and nationality on gaming experience, both variables are of interest as a potential influencing factor. This is because the disparities between the two groups were visible on the box plots. For gender, it might be that visual novels have the stereotype of being dating sims, and therefore could cause a different gaming experience, depending on gender. However, this is a speculation at best. Furthermore, *Diverse Perspectives* does not focus at all on forming romantic relationships, so it is unclear yet what role gender exactly plays.

Regarding nationality, because *Diverse Perspectives* focused on being a new student looking for accommodation in a new city, it was especially relatable to international students.

This is because they often experienced the same struggles, while not all Dutch students need to look for a room when they are studying in Utrecht and might already have built a social network from their previous studies. These discrepancies in life experience could have caused a difference in gaming experience as well. We encourage further research to validate and analyse these exploratory results.

7.1.1.2 Players' choices

Regarding the Sankey Chart, it is remarkable to see that most players chose the same options, as the alternative branches are often very thin. This is also reflected in the choice distributions table. Overall, the participants seemed to have a strong moral compass. As stated before, while this may not be the target group that needs this game the most, they are likely more receptive to it than people with very strong conservative ideas.

Most participants said they chose the answer that fitted their own judgement, instead of choosing the morally 'best' answer. Additionally, various participants mentioned that just thinking and reflecting about these dilemmas already taught them something. We perceive this as a positive development. Before the experiment started, we hypothesised that every participant would just play to reach the 'best' ending and the explicit educational value would then lie in the recap at the end of the game. However, in hindsight, it is more educational that people are thinking critically about their own decisions and pick up some skills along the way, instead of merely wanting to achieve the highest score.

7.1.2 Qualitative results

The results are interpreted by describing how the five main categories could have had an impact on the forming of the preliminary theories that were created from the selective coding. Furthermore, we discuss what their implications could be for further research. An explanation of what each (child) code represents, can be found back in section 6.3.1.

The first theory (Would vs Should) was based on the *Dilemmas* category, as it had the child codes 'Would vs Should', 'Real life' and 'Exploration'. The reason why players experienced this dilemma, could be explained by the fact they knew what the 'best' option was due to their morals, however this was inconsistent with the behaviour that they actually performed in daily life. This conflict could cause a dilemma, because people prefer to be consistent in their behaviour (Cialdini & Goldstein, 2004). This could also explain why most players chose to stick to the options in real life. The few players that chose to explore, could have done so because the game itself is perceived as a safe environment (Froschauer et al., 2010). This is because the in-game consequences do not affect the real world. It would be valuable to further investigate which factors could play a role in this Would-versus-Should dilemma. For instance, the BrainHex player type (Nacke et al., 2014). More about BrainHex can be found in section 7.3.3.

The second theory (Strangers vs Friends) was again based on the *Dilemmas* category and the *Visuals* as well, with its (grand)child codes 'Strangers vs Friends' and 'Characters'. Characters appear as strangers, when players see them for the first time. They can be promoted to friends when they recur in the game, so that the player is able to form a bond with them. Uncertainty Reduction Theory could explain why players are more comfortable confronting

friends. Since we form a mutual bond with friends, we are better at anticipating their behaviour. Thus, they are less likely to violate our expectations, which in turn reduces our unpleasant experience of uncertainty. Hence, we could feel more confident in confronting our friends, as we have an idea of how the other will react (Berger & Calabrese, 1974). For further research, it would be worthwhile to investigate how players behave when their friends are involved in a dilemma.

Theory three (Trivial choices are not trivial) is linked to the *Dilemmas* and *Narrative* categories, since it has the child codes 'Trival vs Consequential' and 'Customisation'. Since making moral choices is a cognitive demanding task, players perceived it as nice to have a break from it to lower the cognitive load. Moreover, the trivial choices provided the player some autonomy in shaping their character, which might increase the player's intrinsic motivation to continue playing (Schneider et al., 2018). Hence, we recommended to include these trivial choices in future designs.

The fourth theory (Increase emotional impact) was based on both the *Narrative* and *UX* categories, because it had the child codes 'Identification', 'Length', 'More episodes' and 'Introductory'. The first three codes had children, and they were also used. Since emotions influence how people perceive the world, it impacts the player's judgement and decision-making as well. Emotions are even able to cast a bias over decisions. Nevertheless, they might also be helpful in steering social decisions, since one's own emotions could help understand the emotions of another (Lerner et al., 2015). It would be interesting if future works included emotionally loaded situations. For example, characters showing explicit anger and fear both verbally and nonverbally, and researched their potential effect on the player's decisions.

The *Dilemmas* category impacted the fifth theory (Perspective switch stimulates reflection), due to its child code 'Perspective switch' and its children. As switching perspective elicits cognitive empathy, it could make people more understanding of the other's position. Moreover, by hearing the motivation behind the NPC's actions, it allows players to discover potential similarities between them and not only differences (Belman & Flanagan, 2010). Future studies could extend this perspective switch. For instance, allowing the avatar to have a full dialogue with the NPC that they are switching with, while reading their mind.

Theory six (Raising awareness) is linked to the *UX* category. The children of the child code 'Usefulness' were used. It is difficult to change the mind of extreme people, as it is likely that their level of reception is low, especially when the game is played on a voluntary basis. Furthermore, a change in attitude does not necessarily mean a change in behaviour (Noon, 2018). Because the material is encapsulated in a game, it might be more beneficial to younger players as well, since it provides an informal and active way of engaging with the content. Additionally, prejudices and biases are already formed at a young age, hence highlighting the relevance of *Diverse Perspectives* to alleviate the harmful effects of these processes (Brereton & Young, 2022). Future work could research what design choices work best with the needs and preferences of younger audiences.

The seventh theory (Relevance) was based on the *Narrative* and the *Visuals* categories. The child codes and their children were 'Dialogue flow', 'Identification' and 'Familiar background'. According to ARCS, relevance is one of the four requirements to maintain motivation to continue learning. Furthermore, *Diverse Perspectives* contains many trivia facts about Utrecht.

This is to obtain the attention of not only people who are new to Utrecht. Attention is part of the ARCS requirements as well (Keller, 1987; Li & Keller, 2018). To study the influence of the visuals and facts on relevance, alternative versions of the game could be made that occur in different cities, but with similar narratives. The new target group would then be the students of that particular city.

Theory eight (Replay) was impacted by the *Dilemmas* and *UX* categories, as they contain the 'Exploration' and 'Endings' child codes. For players, it can be rewarding to discover all paths. Likewise, *Diverse Perspectives* presents a sort of moral mirror to the player. This is because players tend to perceive their ending as a reflection on their moral behaviour. According to the *Three Domains of Self* from Higgins (1987), it is plausible that the ending people achieved, does not comply with their 'ideal' or 'ought' self. In other words, they keep trying until they are content with what they see in the mirror. It would be interesting to further investigate the effects of replay and if it has any connections to the Three Domains of Self.

In sum, these theories could be useful as inspiration for further research. Especially the theories 'Would vs Should' (theory 1), 'Strangers vs Friends' (theory 2) and 'Perspective switch stimulates reflection' (theory 5) seem of high relevance to be investigated more thoroughly, since *Diverse Perspectives* is in essence a social dilemma game.

7.1.2.1 Implications of the preliminary theories

As these preliminary theories appear to have strong links to several psychological theories, it seems to be fruitful to increase the focus on the psychological aspects of the game, besides the pedagogical aspects. For instance, the concept of group dynamics could be further investigated. This is because a lot of dilemmas happen in a social context and the relationships between the avatar and the NPCs affect the player's decisions. Additionally, the *Dilemma Game* is played in a group setting. This offers players the opportunity to discuss scenarios with each other and reflect on their own opinions as well as those of others. Therefore, it seems worthwhile to create a multiplayer variant of *Diverse Perspectives*, to further research the effect of group dynamics and how groups are formed in the first place.

Moreover, another area of interest could be the personality of the player. This is because *Diverse Perspectives* requires the use of several cognitive functions, such as emotion regulation, reflection and decision-making. These psychological processes are unique to each player and could serve as an useful source for personalisation of the game to better suit their needs and preferences. For instance, the intention of the perspective switch is to invoke empathy and compassion for the other. However, the degree of empathy differs between players, as this trait is related to their personality. Thus, it is highly likely the player's personality influences the effectiveness of the perspective switch.

Furthermore, the link between attitude and behaviour could be more thoroughly researched. While a change in the first does not necessarily result in a change in the latter, a positive change in attitude seems to be a promising starting point. *Diverse Perspectives* allows players to actively choose how their avatar would behave. By giving them the power to shape the narrative, it intents to make the link between attitude and behaviour more explicit. The goal of *Diverse Perspectives* is to encourage diversity. This can be reached by inclusive behaviour, which in turn relies on inclusive mindsets.

In other words, *Diverse Perspectives* aims to present the player with a 'moral mirror' and intends to equip them with the tools to achieve what they would like to see in that mirror.

7.2 Limitations

The research encountered some limitations that could have affected the results. Suggestions on how these limitations could be mitigated are described below.

7.2.1 Participant sample

First, the small sample size of the participants and the use of non-parametric tests prevents generalisation of the data. This is especially the case with unequal sample sizes. It is hard to say something about the reliability of the results, since a single data point could have a disproportionate big influence. As the design had an iterative approach, every step involved the recruitment of new participants. Requirements for participants were that they studied at the UU and that they could participate on campus. 16 participants completed both parts per condition. As the experiment had a between-subjects design, this results in 32 participants in total. Ideally, there would be at least 30 participants per condition to be able to detect any significance, which means a minimum of 60 participants overall.

Second, three participants attended the colloquium and another three participants were in the researcher's group from the *Serious Gaming* course last year. This implies that they already had some prior knowledge about the experiment, which could have influenced their results. In a similar vein, almost half of the participants had a HCI or Game and Media Technology background. Hence, they might already be a bit biassed, as they are familiar in this field of study. In follow-up studies, a larger and more diverse participant pool should be used to alleviate this limitation.

Moreover, the experiment did not take into account that some participants could have dyslexia. As it required a lot of reading, the trial for these participants durated longer than intended and this could have affected their results. Further research should control for this variable.

Lastly, P28 did not seem to have fully understood the experiment. Evidence why we suspect this, consists of the participant asking mid-game when the game would start, while they had received explicit instructions that the game would begin. In the interview, P28 commented that they saw *Diverse Perspectives* as an extended part of the survey, because of the dilemmas. Moreover, they would like to have a clearer goal in the game, more feedback after they made a choice, and a clarification of the educational part. It is important to note that this participant was in the control condition, so they missed all the perspective switches. This could explain why *Diverse Perspectives* was not perceived as educational as intended.

As their constructive feedback was still regarded as useful, their data was not treated as an outlier. However, their results should be interpreted with caution, since it is likely this participant might have misunderstood the goal of the experiment.

7.2.2 Risk of socially desired answers

The M-GUDS-S measures one's attitude towards a relatively loaded topic and the game appeals to one's morals as well. Hence, there exists the risk of socially desirable answers, in both the game and the questionnaire. This could be because participants might be afraid

to be cast away as being racist or narrow-minded. While the data has been anonymised, participants had to fill in their participant ID to connect their pre-test results with their post-test results. The existence of this ID might have let participants feel less anonymous in their answers. Although participants could explain and reflect about their choices in the interview, there still exists a chance that participants felt their actions were being 'graded'. To soothe this risk, future studies should stress that their participant ID is merely used to link their pre-test and post-test data and that their participation is not judged in any way at all. That is, their participation is not a test, but rather delivers valuable contributions to the research.

Likewise, questionnaires always have a degree of subjectiveness inherent to them, as they are based on self-reports. Even if the M-GUDS-S showed differences after playing the game, a change in attitude does not necessarily mean a change in behaviour as well. Furthermore, the M-GUDS-S seems somewhat outdated, as it originates from 2000 (Fuertes et al., 2000). Participants were confused by some questions and offered suggestions. As the M-GUDS-S is a validated research tool, we decided not to change the questionnaire. However, it might be time for a revised version that better suits the contemporary world.

7.2.3 Risk of researcher's bias

Diverse Perspectives is only made by the researcher. Even though we adopted a user-centred and iterative approach, the game still represents the researcher's moral views, which is a subjective representation. Another point of critique is that the NPCs were still quite one-dimensional, as their dilemma only represented one aspect of diversity. In reality, people could face multiple marginalising factors, referring back to the concept of intersectionality in section 2.4.1. This was not represented in the game, because the intention was to let the player focus on just one dilemma at a time. Yet, this simplification of reality does not represent the intersectional issues people might face in their daily lives.

Moreover, the high-fidelity prototype is not scientifically validated, as it is only used in this study. Future studies could use the prototype to investigate its validity and reliability. Likewise, collaboration with pedagogical and diversity experts would increase the educational value of the game, as the researcher has little expertise in both fields.

In a similar vein, it is difficult to evaluate the results of the qualitative analysis, as it is a subjective method. While Straussian Grounded theory provides a systematic and interactive approach of data analysis, the researcher may be biassed when interpreting the results. To reduce this risk, multiple coders should be used for the qualitative analysis and their inter-coder reliability should be measured.

7.2.4 The follow-up survey

Every participant was told that they received the follow-up survey a week after the experiment and that they should fill it in to complete their participation. This time span is fairly short, as the link between the survey and the game could still be present in the participant's memory. This could have influenced their results. Due to the current timeframe, it was not feasible to make this time span longer. However, it is interesting to increase this span to a month to make the link between the game and the survey less obvious. Nonetheless, the time span should not be too long, as this would make it harder to test the effects of the game.

7.3 Further research

7.3.1 Actual distribution of the game

The study created various ideas for further research. For example, *Diverse Perspectives* could be distributed at the introduction week, so that all new UU students receive the opportunity to play it. As the game mechanics remain the same and only the narrative needs to be adapted, the game has potential to be transferred to other target groups too, such as students in other cities or students at various educational levels. Additionally, an app version could be developed as well. To be playable anytime and anywhere, it is more convenient to also have a mobile version of *Diverse Perspectives*, instead of only a desktop version. This is because most people have their phone with them and they do not need to boot their laptop to just play the game.

Although it was out of the scope of this thesis, it would be interesting to investigate the effect of replays on the player's choices, gaming experience and attitude towards diversity. As a reward for complete participation, participants did receive a copy of the game to play whenever they like, but no data was collected from it.

7.3.1.1 Potentially other target groups

The current research focused on UU students, as the narrative was based in Utrecht. This target group has a few things in common, such as being in one's young adulthood and studying an university level degree in Utrecht. This specific group was chosen to scope the research, yet, *Diverse Perspectives* has potential to focus on other groups too.

First, the narrative of *Diverse Perspectives* could be translated to other cities, so that its story becomes relevant to the students there. As mentioned before, the issue of negative intergroup biases and exclusion is not only present in Utrecht, but applies to a wider scope as well.

Second, *Diverse Perspectives* could be aimed at a younger target group too. For instance, children that are in middle or high school. At that age, it is likely they encountered others that differ from themselves on several aspects. While their environment and upbringing probably still have a heavy impact on their beliefs, their age is possibly mature enough to allow them to shape their own opinions on diversity matters.

Third, another target group could consist of elderly people. It is likely that a lot of diversity topics are 'new' to them, since those were simply not acknowledged in the past. *Diverse Perspectives* could aid this group to familiarise themselves with these concepts in an informal and playful way, so that this target group would be interested in educating themselves.

Lastly, another group of interest would be people with extreme beliefs. It seems worthwhile to research where the roots of their beliefs stem from and if they can be challenged. Furthermore, their vision on these dilemmas could taken into account in the story as well, to increase its resemblance to reality.

The key thing that these groups should have in common, is being open to familiarise themselves with these topics of diversity and not be afraid to encounter different opinions.

7.3.2 Game improvements

To increase the immersion of *Diverse Perspectives*, the questionnaires could be integrated into it. For instance, some of the player's demographic information could be linked to their avatar, such as their gender and nationality, so that the character they are playing feels more personalised. This could increase player identification. In a similar vein, statements that were presented in the M-GUDS-S could be represented as scenarios in the game. Additionally, the game could have some lore, such as optional readings, to gather more information on a certain topic. This additional content could guide the player in making decisions (Øygardslia et al., 2020). The player can read lore, so that they unlock new dialogue options when talking to NPCs. For instance, players could interact with objects such as a book or their phone to allow some exploration in the game world. Moreover, this also brings some variety in how the story progresses.

Furthermore, the qualitative data contained a lot of suggestions for improvements. These are categorised in a hierarchical tree structure in Figure 6.12e. These suggestions are divided into three subcategories: 1) Dilemmas, 2) Gameplay, and 3) Visuals.

For the *Dilemmas*, participants indicated that the dilemmas could be more complicated. This could be done by creating more subchoices and increasing the granularity, which draws the player deeper into the situation and allows more nuance. Likewise, more subtle and less common dilemmas could be represented, so that players are really challenged to reflect on their behaviour. Alternatively, the game could present the player with some dilemmas, but put them in an observer role, so they do not have to actively intervene every time. A few participants mentioned that the narrative could pick a subset of the most common dilemmas instead and expand on them more, as they seemed to find the amount of topics quite large.

Regarding the *Gameplay*, a few participants said it would be nice if the game had a little tutorial on how the controls worked. While they were seen as intuitive and easy, the researcher did explain the controls before playing. Other tips included a 'Pause' and a 'Save' button. Because *Diverse Perspectives* does not use any timing mechanics, a pause button was not present in the high-fidelity prototype. Nevertheless, being able to pause the game would be nice. Besides that the game automatically saved the player's progress, it was also possible to manually save the game in slots. However, this was intentionally not explained to the players. Since they would start and finish the game in one session, there was no need to manually save their progress during the session.

Moreover, some participants remarked that the goal of the game could be made more clear. While the avatar's goal is explained in the beginning of the story, it could be repeated more often. For example, in the dialogue or in the player's actions. Other feedback consisted of increasing the interactivity of the game with the inclusion of minigames or free exploration of the world, as *Diverse Perspectives* has currently a point-and-click format. Furthermore, several participants talked about having multiple versions of the game, so that players could choose one, depending on if they wanted to explore or really learn something. Additionally, participants stated that they would have liked to personalise their avatar more. For instance, they could indicate whether or not they are already living in Utrecht. Lastly, *Diverse Perspectives* could have sounds, such as people talking or the tram passing by, as it could increase immersion.

Looking at the *Visuals*, some participants mentioned that the characters were a bit awkwardly

cropped off, since only their heads were visible. As explained in section 4.2.4, this was to crop the watermark off. A solution to this could be to use another style, for example a cartoonish one. Other participants said that they would like to have a progress bar, as this could indicate at which point they are in the narrative. Participants also remarked that the perspective switch could have been made more clear visually, such as a special animation or some additional graphics. At last, several participants expressed that the characters' speech and the Utrecht trivia could be distinguished more, as it was sometimes not clear who was saying what. For example, the facts could be interweaved more naturally in the story or they could be shown in a visually different text box.

7.3.3 Player diversity

Players differ from each other too and this could have influenced the results. It would be interesting to take these characteristics into account. Two aspects that seem fruitful are player type and culture.

While there exist several player categorisation scales, BrainHex (Nacke et al., 2014) seems the most fitting. It is empirically based and validated with a large number of participants. The model consists of seven player types: 1) survivor, who likes to escape risks, 2) daredevil, who likes moving around at high speeds while still being in control, 3) conqueror, who likes to struggle until victory is achieved, 4) socialiser, who likes to help others, 5) seeker, who likes to explore, 6) achiever, who likes to collect and achieve everything, and 7) mastermind, who likes solving puzzles and creating strategies.

Players have a primary type (their dominant type) and a secondary type (their subtype), so these types are not mutually exclusive. Their type represents a strong link to their preferred game elements and player needs. Player types may play a beneficial role in designing games, as they provide insights into the needs and wishes of players (Nacke et al., 2014).

Besides different player types, cultures between players could differ as well. Hofstede (1986, 2001) devised five major cultural dimensions on which cultures are different from another. These are: 1) power-distance, 2) individualism-collectivism, 3) masculinity-femininity, 4) uncertainty avoidance, and 5) long-term orientation. Differences between the cultural dimensions of the player's culture and Dutch culture could influence their gaming experience as well.

7.4 Main findings of the Discussion

As stated before, there was not enough evidence to confirm our hypotheses. Some findings even seemed to support the opposite idea of what was hypothesised. Nonetheless, we found three significant results.

First, post-test DC scored higher than pre-test DC. This could imply that *Diverse Perspectives* did have a positive influence on the post-test score, at least regarding this specific component. Second, the scores for Immersion in the marginalised group were lower than in the non-marginalised group. This could be explained by the fact that the scenarios could trigger the player's memories of similar events and hence diminishing the immersion. Third, the scores for Positive Affect in the marginalised group were lower than in the non-marginalised group. This could be caused by the fact that the scenarios could invoke some mixed feelings based on the player's own experiences, and thereby decreasing the positive affect.

A reason why the perspective switch did not have any significant influence on the change in one's attitude towards diversity and on one's gaming experience, could be that the switch was simply too short and too infrequent to create such an impact.

Although the exploratory results did not yield any significant findings, they still seem to be fruitful to investigate in further research.

Next, we elaborated upon each preliminary theory by analysing its possible connections to the five main categories and by linking it to existing theories in the literature. While each theory could be promising to use in future works, three theories appear to be especially interesting. These are: 'Would vs Should' (theory 1), 'Strangers vs Friends' (theory 2) and 'Perspective switch stimulates reflection' (theory 5). This is because *Diverse Perspectives* is in essence a social dilemma game. Likewise, the possible implications of these theories were discussed.

The research encountered some limitations as well. Among these are the small participant sample size (N = 32), the risk of socially desired answers, the risk of the researcher's bias and the one-week time span between part 1 and part 2 of the experiment.

Lastly, ideas for further research were offered. For instance, creating a mobile version and multiple variations of *Diverse Perspectives* were suggested for when the game would be actually distributed. Likewise, potentially broader target groups were discussed as well. Moreover, various suggestions for improvements were proposed. Examples are increased personalisation of the game by linking the player's demographics to their avatar, such as their BrainHex player type and culture. Other recommendations included adding lore and a tutorial, increasing the complexity of the dilemmas, clarifying the goal of *Diverse Perspectives* and improving the visuals.

Chapter 8

Conclusion

The aim of this research was to encourage inclusive behaviour among UU students by stimulating interactions between diverse groups. This was done by creating a serious game in the form of a visual novel, called *Diverse Perspectives*. The target users were UU students. In the field of serious games, a gap was identified on the combination of diversity with (international) student inclusion. Hence, this research intended to contribute to the field of inclusive serious games. The research question this thesis aimed to answer was: *How could Diverse Perspectives be used to promote diversity at Utrecht University?* The research question was divided into three subquestions, namely:

- *R1: What is the impact of the perspective switch on one's attitude towards diversity?*
- *R2: Which factors contribute to a positive gaming experience?*
- *R3: What is the effect of identifying with a marginalised group on the gaming experience?*

We reviewed related work to get a grasp of the concepts that were fundamental in creating the serious game, such as storytelling, pedagogical theories and diversity. Next, *Diverse Perspectives* was developed by following a user-centred and iterative approach. That is, the target users were involved early in the design process. Moreover, a low-fidelity prototype was made in Figma and evaluated. This feedback was incorporated when creating the high-fidelity prototype in Ren'Py. Then, this high-fidelity prototype was used in the experiment with a mixed-method and between-subjects design, where we measured the participant's attitude towards diversity with the M-GUDS-S questionnaire and their gaming experience with the GEQ questionnaire. The M-GUDS-S questionnaire served to assess whether the objective of promoting diversity through serious games could be reached. The GEQ questionnaire was used to evaluate the performance of the serious game itself and if it had potential to be an effective intervention tool.

Regarding R1, participants indicated that they found the perspective switch helpful as it creates understanding for the other person's view. Yet, we only found a significant result in the change of DC, but not for RA and CD. Concerning R2, four factors have been identified. These are: 1) Immersion, 2) Flow, 3) Competence, and 4) Positive Affect. Furthermore, the qualitative data revealed that the depth and relatability of the narrative play a vital role, as well as the dialogue flow and navigation clarity. Looking at R3, we found that identification

with a marginalised group had a big effect on Immersion and Positive Affect, as both scored significantly lower compared to people who did not identify with a marginalised group. For the first group, Flow scored lower as well, while Competence scored higher. However, these differences were not significant. Although the exploratory research yielded no significant results, they are still promising to be investigated more thoroughly in the future.

In sum, while our hypotheses could not be confirmed from the quantitative results, the qualitative results indicated that there is potential for *Diverse Perspectives* in further research. To analyse the qualitative data, we followed the Straussian Grounded Theory approach that combined a priori coding with emergent coding. Consequently, a hierarchical coding tree was developed that consisted of five main codes, which are: 1) Narrative, 2) Dilemmas, 3) Visuals, 4) UX, and 5) Improvements. These codes allowed the construction of eight preliminary theories. Especially the theories 'Would vs Should' (theory 1), 'Strangers vs Friends' (theory 2) and 'Perspective switch stimulates reflection' (theory 5) seem to be of high relevance, since *Diverse Perspectives* is in essence a social dilemma game. We encourage future works to investigate and validate these theories, as well as their implications. Hence, the gap between serious games, diversity and (international) student inclusion remains to be filled.

References

Allen, K.-A., Gray, D. L., Baumeister, R. F., & Leary, M. R. (2022). The need to belong: A deep dive into the origins, implications, and future of a foundational construct. *Educational Psychology Review*, *34*(2), 1133–1156.

Backx, S., Chen, S., Mantz, T., Peters Rit, G., & Vleeshouwers, C. (2022). Change Perspective: A Serious Game For Cultural Integration Between International And Dutch Students. [Project Report]. *INFOMSEGA course, Utrecht University.*

Barrera Yañez, A. G., Alonso-Fernandez, C., & Fernandez Manjon, B. (2020). Review of serious games to educate on gender equality. *Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality*, 662–668.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Interpersonal development*, 57–89.

Belman, J., & Flanagan, M. (2010). Designing games to foster empathy. *International Journal of Cognitive Technology*, *15*(1), 11.

Berger, C. R., & Calabrese, R. J. (1974). Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. *Human communication research*, *1*(2), 99–112.

Blandford, A., Furniss, D., & Makri, S. (2016). Qualitative hci research: Going behind the scenes. *Synthesis lectures on human-centered informatics*, *9*(1), 1–115.

Bouchard, R. Philip. (2017). *How i managed to design the most successful educational computer game of all time.* Retrieved December 20, 2022, from https://medium.com/the-philipendium/how-i-managed-to-design-the-most-successful-educational-computer-game-of-all-time-4626ea09e184

Braidotti, R. (2011). Sexual difference theory. In *Nomadic subjects: Embodiment and sexual difference in contemporary feminist theory* (2nd ed., pp. 91–115). Columbia University Press.

Brereton, J. S., & Young, K. (2022). Establishing social learning in an engineering mooc: Benefits for diversity and inclusion in engineering education. *Sustainability*, *14*(9), 1–24.

Byrne, D. (1997). An overview (and underview) of research and theory within the attraction paradigm. *Journal of Social and Personal Relationships*, *14*(3), 417–431.

Carstensdottir, E., Kleinman, E., & El-Nasr, M. S. (2019). Player interaction in narrative games: Structure and narrative progression mechanics. *Proceedings of the 14th International Conference on the Foundations of Digital Games*, 1–9.

Chabebe Rivera, A. M., Wu, W., & Xue, C. (2022). An augmented reality update of a classic game: "where in the world is carmen sandiego?", case study. *HCI in Games: 4th International Conference, HCI-Games 2022, Held as Part of the 24th HCI International Conference, HCII 2022, Virtual Event, June 26–July 1, 2022, Proceedings*, 569–585.

Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, *55*, 591–621.

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative sociology*, *13*(1), 3–21.

Crenshaw, K. W. (1988). Race, reform, and retrenchment: Transformation and legitimation in antidiscrimination law. *Harvard Law Review*, *101*(7), 1331–1387.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. Harper; Row.

Deci, E. L., & Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian psychology*, *49*(3), 182.

Dörner, R., Göbel, S., Effelsberg, W., & Wiemeyer, J. (2016). Introduction. In *Serious games* (pp. 1–34). Springer.

Dovidio, J. F., Love, A., Schellhaas, F. M., & Hewstone, M. (2017). Reducing intergroup bias through intergroup contact: Twenty years of progress and future directions. *Group Processes & Intergroup Relations*, *20*(5), 606–620.

Dwyer, H., & Smith, J. (2020). A mandatory diversity workshop for faculty: Does it work? *To Improve the Academy*, *39*(2), 183–212.

Elsherif, M. M., Middleton, S. L., Phan, J. M., Azevedo, F., Iley, B. J., Grose-Hodge, M., Tyler, S. L., Kapp, S. K., Gourdon-Kanhukamwe, A., Grafton-Clarke, D., & et al. (2022). Bridging neurodiversity and open scholarship: How shared values can guide best practices for research integrity, social justice, and principled education. *[Unpublished manuscript]*. https://doi.org/10.31222/osf.io/k7a9p

Equality, Diversity & Inclusion (EDI) Office of Utrecht University. (2022). *The Dilemma Game*. Retrieved December 4, 2022, from https://www.uu.nl/en/dilemma-game

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using r*. Sage publications.

Fonseca, X., Slingerland, G., Lukosch, S., & Brazier, F. (2021). Designing for meaningful social interaction in digital serious games. *Entertainment Computing*, *36*, 100385.

Froschauer, J., Seidel, I., Gärtner, M., Berger, H., & Merkl, D. (2010). Design and evaluation of a serious game for immersive cultural training. *2010 16th international conference on virtual systems and multimedia*, 253–260.

Fuertes, J. N., Miville, M. L., Mohr, J. J., Sedlacek, W. E., & Gretchen, D. (2000). Factor structure and short form of the miville-guzman universality-diversity scale. *Measurement and Evaluation in Counseling and Development*, *33*(3), 157–169.

Göbel, S., & Wendel, V. (2016). Personalization and adaptation. In R. Dörner, S. Göbel, W. Effelsberg, & J. Wiemeyer (Eds.), *Serious games* (pp. 161–210). Springer.

Gómez-Zará, D., Guo, M., DeChurch, L. A., & Contractor, N. (2020). The impact of displaying diversity information on the formation of self-assembling teams. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–15.

Harrison-Bernard, L. M., Augustus-Wallace, A. C., Souza-Smith, F. M., Tsien, F., Casey, G. P., & Gunaldo, T. P. (2020). Knowledge gains in a professional development workshop on diversity, equity, inclusion, and implicit bias in academia. *Advances in Physiology Education*, *44*(3), 286–294.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological review*, *94*(3), 319–340.

Hofstede, G. (1986). Cultural differences in teaching and learning. *International Journal of intercultural relations*, *10*(3), 301–320.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations* (2nd ed.). Thousand Oaks: Sage Publications.

Homan, A. C. (2019). Dealing with diversity in workgroups: Preventing problems and promoting potential. *Social and Personality Psychology Compass*, *13*(5), 1–15.

hooks, B. (1991). Theory as liberatory practice. *Yale Journal of Law and Feminism*, *4*, 1–12.

IJsselsteijn, W. A., Kort, Y. D., & Poels, K. (2013). The Game Experience Questionnaire. [Questionnaire]. *Eindhoven University of Technology.*

Johnson, D., Gardner, M. J., & Perry, R. (2018). Validation of two game experience scales: The player experience of need satisfaction (pens) and game experience questionnaire (geq). *International Journal of Human-Computer Studies*, *118*, 38–46.

Joshi, A., Kale, S., Chandel, S., & Pal, D. K. (2015). Likert scale: Explored and explained. *British journal of applied science & technology*, *7*(4), 396–403.

Keller, J. M. (1987). Development and use of the arcs model of instructional design. *Journal of instructional development*, *10*(3), 2–10.

Kottke, J. L. (2011). Additional evidence for the short form of the universality-diversity scale. *Personality and Individual Differences*, *50*(4), 464–469.

Krath, J., Schürmann, L., & Von Korflesch, H. F. (2021). Revealing the theoretical basis of gamification: A systematic review and analysis of theory in research on gamification, serious games and game-based learning. *Computers in Human Behavior*, *125*, 106963.

Kruglanski, A. W., Fishbach, A., Woolley, K., Bélanger, J. J., Chernikova, M., Molinario, E., & Pierro, A. (2018). A structural model of intrinsic motivation: On the psychology of means-ends fusion. *Psychological Review*, *125*(2), 165.

Kukulska-Hulme, A., Gaved, M., Paletta, L., Scanlon, E., Jones, A., & Brasher, A. (2015). Mobile incidental learning to support the inclusion of recent immigrants. *Ubiquitous Learning: an international journal*, *7*(2), 9–21.

Laamarti, F., Eid, M., & Saddik, A. E. (2014). An overview of serious games. *International Journal of Computer Games Technology*, 1–15.

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annual review of psychology*, *66*, 799–823.

Li, K., & Keller, J. M. (2018). Use of the arcs model in education: A literature review. *Computers & Education*, *122*, 54–62.

Loes, C. N., Culver, K., & Trolian, T. L. (2018). How collaborative learning enhances students' openness to diversity. *The Journal of Higher Education*, *89*(6), 935–960.

Lorde, A. (1984). Age, race, class, and sex: Women redefining difference. In *Sister outsider: Essays and speeches* (pp. 114–123). Crossing Press.

Loveless, Becton. (2022). *Cognitive Load Theory - The Definitive Guide.* Retrieved December 28, 2022, from https://www.educationcorner.com/cognitive-load-theory/

Mildner, P., & Mueller, F. '. (2016). Design of serious games. In R. Dörner, S. Göbel, W. Effelsberg, & J. Wiemeyer (Eds.), *Serious games* (pp. 57–82). Springer.

Moreu, G., Isenberg, N., & Brauer, M. (2021). How to promote diversity and inclusion in educational settings: Behavior change, climate surveys, and effective pro-diversity initiatives. *Frontiers in Education*, *6*, 668250.

Nacke, L. E., Bateman, C., & Mandryk, R. L. (2014). Brainhex: A neurobiological gamer typology survey. *Entertainment computing*, *5*(1), 55–62.

Naul, E., & Liu, M. (2020). Why story matters: A review of narrative in serious games. *Journal of Educational Computing Research*, *58*(3), 687–707.

Nieubuur, B. (2020). *Utrecht* (4th ed.). Achter de Dom.

Nieubuur, B. (2022). *Nog meer utrecht* (1st ed.). Achter de Dom.

Nilsson, P. A. (2019). The buddy programme: Integration and social support for international students. *Journal of Comparative & International Higher Education*, *11*, 36–43.

Noon, M. (2018). Pointless diversity training: Unconscious bias, new racism and agency. *Work, employment and society*, *32*(1), 198–209.

Olson, D. M., & Harrell, D. F. (2020). "i don't see color": Characterizing players' racial attitudes and experiences via an anti-bias simulation videogame. *Proceedings of the 15th International Conference on the Foundations of Digital Games*, 1–4.

Øygardslia, K., Weitze, C. L., & Shin, J. (2020). The educational potential of visual novel games: Principles for design. *Replaying Japan*, *2*(2).

Paletta, L., Dunwell, I., Gaved, M., Bobeth, J., Efremidis, S., Luley, P., Kukulska-Hulme, A., Freitas, S. d., Lameras, P., & Deutsch, S. (2013). Advances in maseltov–serious games in a mobile ecology of services for social inclusion and empowerment of recent immigrants. *International Conference on Advances in Computer Entertainment Technology*, 440–455.

Pietri, E. S., Hennes, E. P., Dovidio, J. F., Brescoll, V. L., Bailey, A. H., Moss-Racusin, C. A., & Handelsman, J. (2019). Addressing unintended consequences of gender diversity interventions on women's sense of belonging in stem. *Sex Roles*, *80*(9), 527–547.

Pitz, G. F., & Harren, V. A. (1980). An analysis of career decision making from the point of view of information processing and decision theory. *Journal of Vocational Behavior*, *16*(3), 320–346.

Poels, K., de Kort, Y. A., & IJsselsteijn, W. A. (2007). D3.3: Game experience questionnaire: Development of a self-report measure to assess the psychological impact of digital games.

Puritty, C., Strickland, L. R., Alia, E., Blonder, B., Klein, E., Kohl, M. T., McGee, E., Quintana, M., Ridley, R. E., Tellman, B., et al. (2017). Without inclusion, diversity initiatives may not be enough. *Science*, *357*(6356), 1101–1102.

Rienties, B., Luchoomun, D., & Tempelaar, D. (2014). Academic and social integration of master students: A cross-institutional comparison between dutch and international students. *Innovations in education and teaching international*, *51*(2), 130–141.

Rienties, B., & Tempelaar, D. (2013). The role of cultural dimensions of international and Dutch students on academic and social integration and academic performance in the Netherlands. *International Journal of Intercultural Relations*, *37*(2), 188–201.

Schell, J. (2015). *The art of game design: A book of lenses* (2nd ed.). CRC Press.

Schneider, S., Nebel, S., Beege, M., & Rey, G. D. (2018). The autonomy-enhancing effects of choice on cognitive load, motivation and learning with digital media. *Learning and Instruction*, *58*, 161–172.

Serebrenik, A. (2022). Inaugural lecture on Social Software Engineering. [Booklet]. *Eindhoven University of Technology.*

Shortt, M., Tilak, S., Kuznetcova, I., Martens, B., & Akinkuolie, B. (2021). Gamification in mobile-assisted language learning: A systematic review of duolingo literature from public release of 2012 to early 2020. *Computer Assisted Language Learning*, 1–38.

Smith, R., & Decker, A. (2016). Understanding the impact of qpoc representation in video games. *2016 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*, 1–8.

Streicher, A., & Smeddinck, J. D. (2016). Personalized and adaptive serious games. In R. Dörner & et al. (Eds.), *Entertainment computing and serious games* (pp. 332–377). Springer.

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational psychology review*, *22*, 123–138.

Sweller, J. (2011). Cognitive load theory. In J. Mestre & B. Ross (Eds.), *Psychology of learning and motivation* (pp. 37–76). Academic Press.

Szolin, K., Kuss, D. J., Nuyens, F. M., & Griffiths, M. D. (2022). Exploring the user-avatar relationship in videogames: A systematic review of the proteus effect. *Human–Computer Interaction*, 1–26.

Szolin, K., Kuss, D. J., Nuyens, F. M., & Griffiths, M. D. (2023). "i am the character, the character is me": A thematic analysis of the user-avatar relationship in videogames. *Computers in Human Behavior*, *143*, 1–13.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behaviour. In *Psychology of intergroup relations* (pp. 7–24). Nelson-Hall.

Tancred, N., Vickery, N., Wyeth, P., & Turkay, S. (2018). Player choices, game endings and the design of moral dilemmas in games. *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*, 627–636.

Tompkins, J. E., & Martins, N. (2022). Masculine pleasures as normalized practices: Character design in the video game industry. *Games and Culture*, *17*(3), 399–420.

Trolian, T. L., & Parker III, E. T. (2022). Shaping students' attitudes toward diversity: Do faculty practices and interactions with students matter? *Research in Higher Education*, *63*(5), 849–870.

Verpoorten, D., Poumay, M., & Leclercq, D. (2007). The eight learning events model: A pedagogic conceptual tool supporting diversification of learning methods. *Interactive Learning Environments*, *15*(2), 151–160.

Wang, A. I., & Tahir, R. (2020). The effect of using kahoot! for learning–a literature review. *Computers & Education*, *149*, 1–22.

Wouters, P., & Oostendorp, H. v. (2017). Overview of instructional techniques to facilitate learning and motivation of serious games. In *Instructional techniques to facilitate learning and motivation of serious games* (pp. 1–16). Springer International Publishing.

Appendix A

Figures

A.1 Narrative branching structure

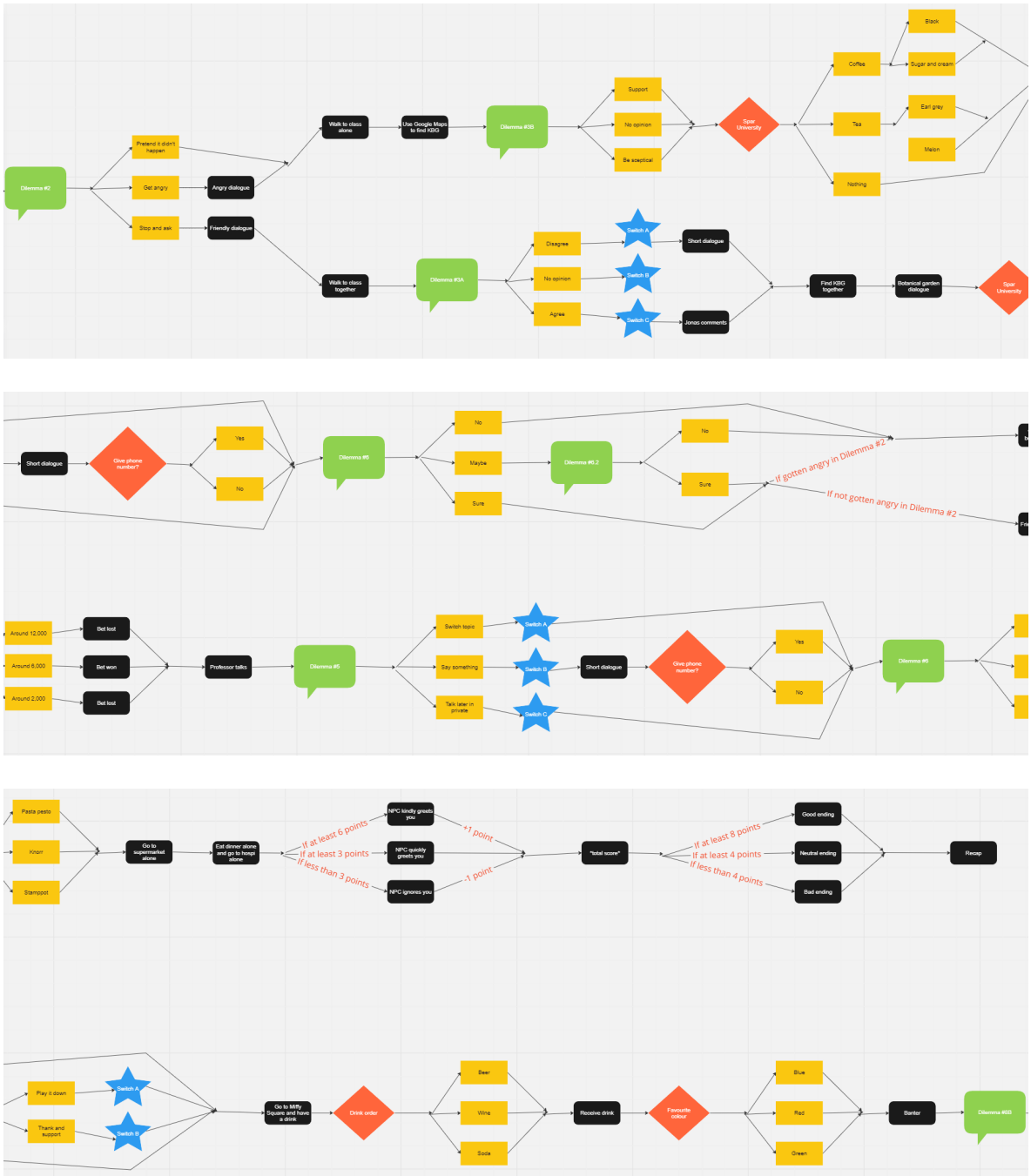


Figure A.1: Snippets of the narrative branches as made in Miro. The arrows go from left to right. Consequential choices are represented by green speech bubbles, while trivial choices are displayed as red diamonds. Orange rectangles show choice options and blue stars denote a perspective switch. Lastly, black boxes depict dialogue scenes.

A.2 Sankey flow chart

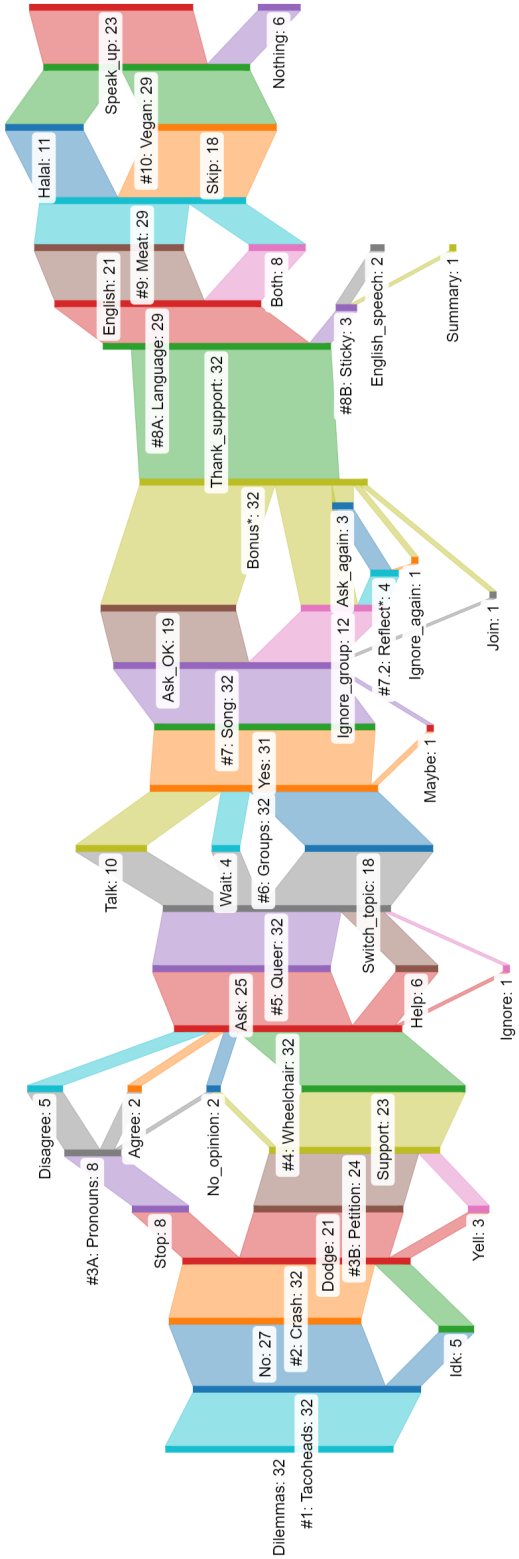
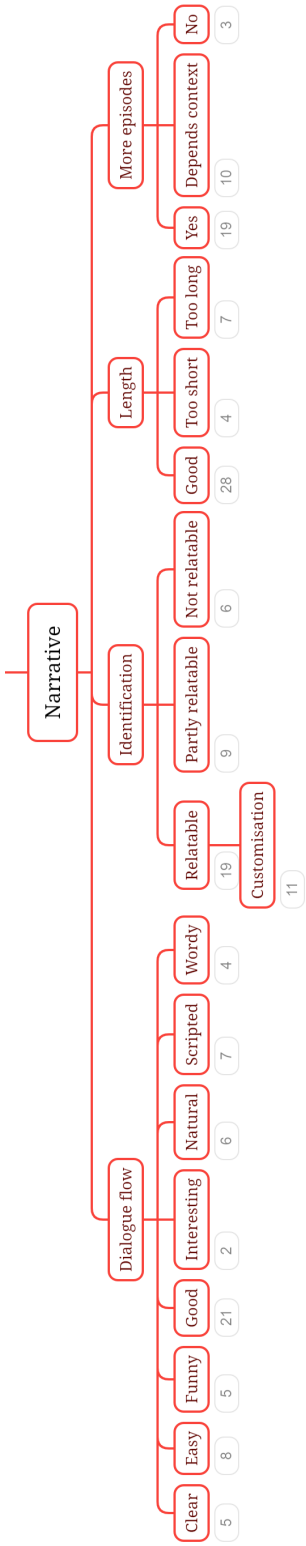
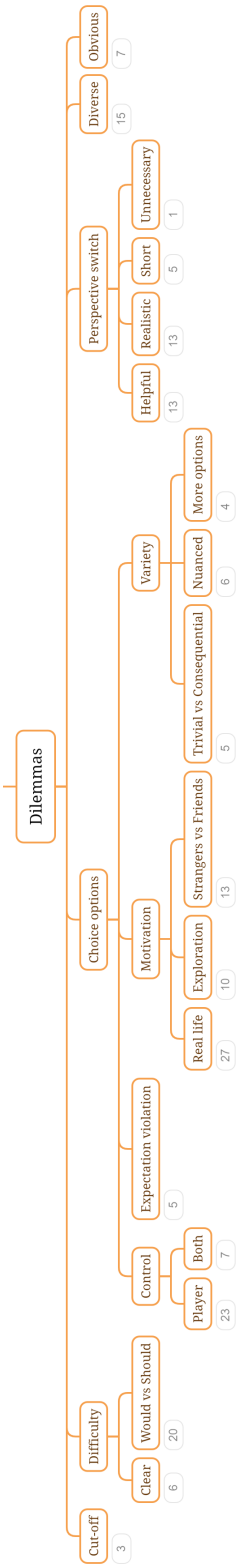


Figure A.2: The Sankey flow chart, which has been made in SankeyMATIC. It shows the choices players made and the dilemmas, with a hashtag (#), they encountered. The number after each label represents the number of players that encountered, where 32 is the maximum number. The width of a branch indicates the number of people that made a certain choice.

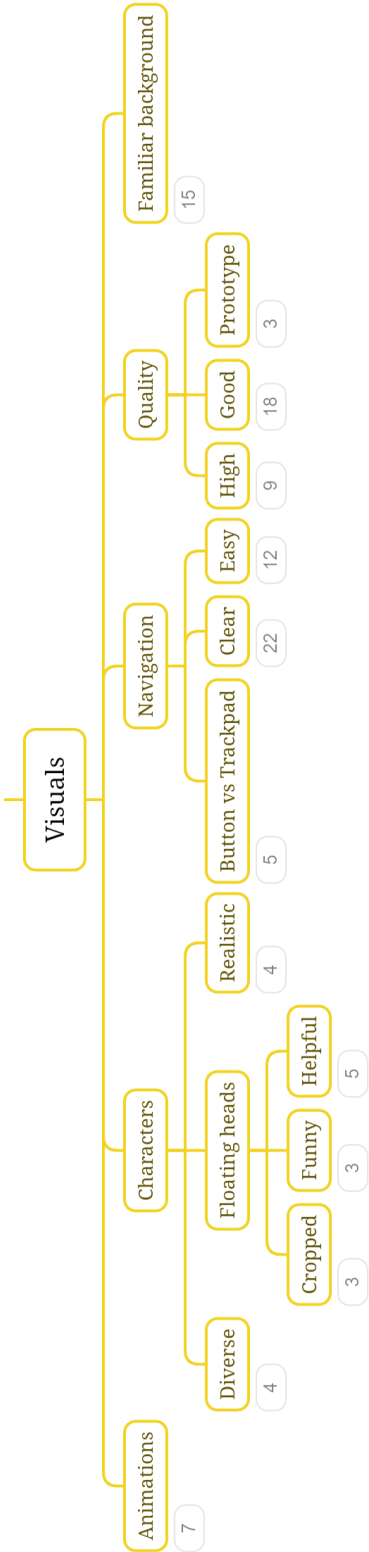
A.3 Hierarchical code structure


Narrative
Dialogue flow
Clear
5
Easy
8
Funny
5
Good
21
Interesting
2
Natural
6
Scripted
7
Wordy
4
Identification
Relatable
19
Customisation
11
Partly relatable
9
Not relatable
6
Length
Good
28
Too short
4
Too long
7
More episodes
Yes
19
Depends context
10
No
3

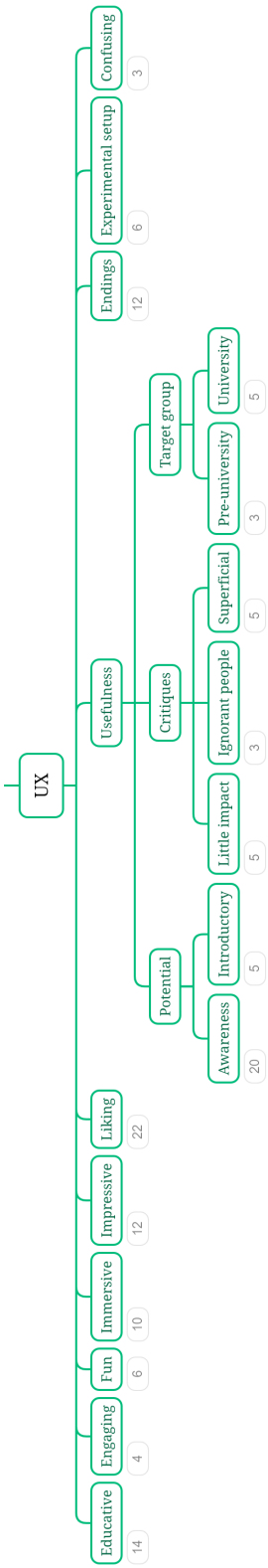
(a) Narrative


Dilemmas
Cut-off
3
Difficulty
Clear
6
Would vs Should
20
Choice options
Control
Player
23
Both
7
Expectation violation
5
Motivation
Real life
27
Exploration
10
Strangers vs Friends
13
Variety
Trivial vs Consequential
5
Nuanced
6
More options
4
Perspective switch
Helpful
13
Realistic
13
Short
5
Unnecessary
1
Diverse
15
Obvious
7

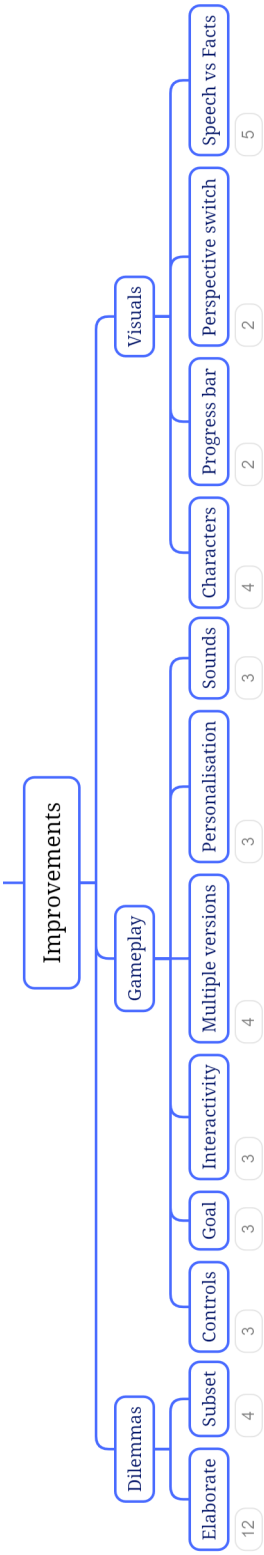
(b) Dilemmas


Visuals
Animations
7
Characters
Diverse
4
Floating heads
Cropped
3
Funny
3
Helpful
5
Realistic
4
Navigation
Button vs Trackpad
5
Clear
22
Easy
12
Quality
High
9
Good
18
Prototype
3
Familiar background
15

(c) Visuals


UX
Educative
14
Engaging
4
Fun
6
Immersive
10
Impressive
12
Liking
22
Usefulness
Potential
Awareness
20
Introductory
5
Critiques
Little impact
5
Ignorant people
3
Superficial
5
Target group
Pre-university
3
University
5
Endings
12
Experimental setup
6
Confusing
3

(d) UX



(e) improvements

Figure A.3: All the children codes of the five parent codes. The number at the bottom of the child nodes indicates the number of participants that mentioned that specific code. The tree diagrams of the hierarchical code structure are made in Xmind.

Appendix B

Tables

B.1 Descriptive statistics M-GUDS-S

Table B.1 shows the descriptive statistics representing the scores of the M-GUDS-S data for the whole sample. Although the medians do not change, the means change slightly. Concretely, the mean for DC has increased a little and the same applies to RA as well. Lastly, the mean for CD has decreased a very small amount. Furthermore, homogeneity of variance between the pre-test and post-test of each component of the M-GUDS-S was assessed with Levene's test. This test yielded no significant results, which means that this assumption was not violated.

Table B.1: The descriptive statistics representing the scores for the M-GUDS-S pre-test and post-test data along the three components for the whole sample.

	Pre DC	**Post DC**	**Pre RA**	**Post RA**	**Pre CD**	**Post CD**
median	25.0	25.0	24.0	24.0	8.5	8.5
mean	23.0	24.7	24.5	24.9	9.4	9.3
SE mean	0.6	0.6	0.5	0.5	0.5	0.6
sd	3.6	3.3	2.6	3.0	2.9	3.2
IQR	3.5	4.5	3.0	3.3	4.3	4.0

Table B.2 shows the descriptive statistics representing the scores of the M-GUDS-S data for version A. Both the medians as well as the means show changes. Concretely, the mean for DC has increased a little and the same applies to RA as well. However, the mean for CD did not change. Furthermore, homogeneity of variance between the groups for each component of the M-GUDS-S was assessed with Levene's test. This test yielded no significant results, which means that this assumption was not violated.

Table B.3 shows the descriptive statistics representing the scores of the M-GUDS-S data for version B. Both the medians as well as the means differ. Concretely, the median and mean for DC has increased a little and the same applies to RA as well. The mean for CD did decrease

Table B.2: The descriptive statistics representing the scores for the M-GUDS-S pre-test and post-test data along the three components for version A (the experimental condition).

	Pre DC_A	**Post DC_A**	**Pre RA_A**	**Post RA_A**	**Pre CD_A**	**Post CD_A**
median	25.0	25.5	24.0	24.5	8.5	8.0
mean	24.9	25.4	23.9	24.3	9.1	9.1
SE mean	0.7	0.5	0.7	0.7	0.6	0.8
sd	2.7	2.2	2.6	2.8	2.5	3.2
IQR	2.3	3.5	2.3	3.0	4.0	1.8

by a very small amount, while the median increased. This indicates a negative skew of the data. Furthermore, homogeneity of variance between the groups for each component of the M-GUDS-S was assessed with Levene's test. This test yielded no significant results, which means that this assumption was not violated.

Table B.3: The descriptive statistics representing the scores for the M-GUDS-S pre-test and post-test data along the three components for version B (the control condition).

	Pre DC_B	**Post DC_B**	**Pre RA_B**	**Post RA_B**	**Pre CD_B**	**Post CD_B**
median	23.0	23.5	24.5	25.0	8.5	10.0
mean	23.0	24.0	25.1	25.4	9.6	9.5
SE mean	1.1	1.0	0.6	0.8	0.8	0.8
sd	4.2	4.1	2.4	3.1	3.3	3.3
IQR	5.5	5.5	3.3	5.5	5.3	4.5

Additionally, Table B.4 shows the descriptive statistics representing the differences (Δ) in the scores of the M-GUDS-S data for both versions. Both the differences in medians as well as the means between DC and RA are visible. Concretely, the mean and median difference for DC is higher in the control condition than in the experimental condition. Regarding the mean, the reverse holds for RA, although the difference is very small. The median does not change. However, there is no difference in the mean and median for DC. Furthermore, homogeneity of variance between the groups for each component of the M-GUDS-S was assessed with Levene's test. This test yielded no significant results, which means that this assumption was not violated.

Table B.4: The descriptive statistics representing the scores for the M-GUDS-S pre-test and post-test data along the three components for the two conditions.

	Δ **DC$_A$**	Δ **DC$_B$**	Δ **RA$_A$**	Δ **RA$_B$**	Δ **CD$_A$**	Δ **CD$_B$**
median	0.5	1.0	0.5	0.5	0.0	0.0
mean	0.6	1.0	0.4	0.3	-0.1	-0.1
SE mean	0.6	0.5	0.7	0.5	0.5	0.6
sd	2.4	2.0	2.7	2.0	1.9	2.5
IQR	3.0	2.0	3.0	2.3	2.0	2.0

B.2 Descriptive statistics GEQ

Table B.5 shows the descriptive statistics representing the scores of the GEQ data for the two conditions. Both the medians as well as the means differ between version A and version B. Concretely, the mean for version A is smaller for all components, except for Competence. Regarding the median, version A scores higher on Immersion and Competence, but lower in Flow and Positive Affect. Furthermore, homogeneity of variance between the groups of the four components of the GEQ was assessed with Levene's test. This test yielded no significant results, which means that this assumption was not violated.

Table B.5: The descriptive statistics representing the scores for the GEQ data along the four components for both conditions.

	Immersion$_A$	**Immersion$_B$**	**Flow$_A$**	**Flow$_B$**	**Competence$_A$**	**Competence$_B$**	**Positive affect$_A$**	**Positive affect$_B$**
median	17.5	17.0	11.0	13.0	14.0	13.5	15.0	16.0
mean	16.8	17.7	12.3	12.4	13.9	12.7	15.6	16.3
SE mean	1.1	0.8	1.1	1.1	0.8	1.1	0.7	0.7
sd	4.4	3.1	4.3	4.3	3.3	4.4	2.9	2.7
IQR	2.8	4.0	6.0	5.5	3.5	3.0	4.3	3.3

Table B.6 shows the descriptive statistics representing the scores of the GEQ data for participants that identified with a marginalised group (Y) and participants who did not (N). Both the medians as well as the means differ between group Y and group N. Concretely, the mean for Y is smaller for all components. Regarding the median, Y only scores higher on Flow. It scores equal on Competence, but lower on Immersion and Positive Affect. Furthermore, homogeneity of variance between the two groups on the four components of the GEQ was assessed with Levene's test. This test yielded no significant results, which means that this

assumption was not violated.

Table B.6: The descriptive statistics representing the scores for the GEQ data along the four components for participants who identify with a marginalised group (Y) and participants who do not (N).

	Immersion$_{\mathbf{Y}}$	**Immersion$_{\mathbf{N}}$**	**Flow$_{\mathbf{Y}}$**	**Flow$_{\mathbf{N}}$**	**Competence$_{\mathbf{Y}}$**	**Competence$_{\mathbf{N}}$**	**Positive affect$_{\mathbf{Y}}$**	**Positive affect$_{\mathbf{N}}$**
median	17.0	18.0	12.0	11.0	14.0	14.0	15.0	16.0
mean	15.0	18.2	11.4	12.4	12.6	13.3	14.7	16.5
SE mean	1.6	0.6	1.3	1.0	1.7	0.7	0.8	0.7
sd	4.9	2.8	4.1	4.5	5.4	3.0	2.5	2.8
IQR	2.8	3.5	7.0	7.5	5.5	3.0	3.0	4.0

B.3 Descriptive statistics M-GUDS-S exploratory

Table B.7 shows the descriptive statistics representing the scores of the M-GUDS-S data for participants that identified with a marginalised group (Y) and participants who did not (N). Although slightly, the means and medians differ between both groups. Furthermore, homogeneity of variance between the groups for each component of the M-GUDS-S was assessed with Levene's test. This test yielded no significant results, which means that this assumption was not violated.

Table B.7: The descriptive statistics representing the scores for the M-GUDS-S pre-test and post-test data along the three components for participants who identify with a marginalised group (Y) and participants who do not (N).

	Δ DC$_{\mathbf{Y}}$	**Δ DC$_{\mathbf{N}}$**	**Δ RA$_{\mathbf{Y}}$**	**Δ RA$_{\mathbf{N}}$**	**Δ CD$_{\mathbf{Y}}$**	**Δ CD$_{\mathbf{N}}$**
median	0.5	1.0	0.5	0.0	0.0	0.0
mean	1.5	0.4	-0.1	0.4	0.3	-0.1
SE mean	0.7	0.5	0.6	0.6	0.7	0.6
sd	2.4	2.1	2.0	2.5	2.1	2.4
IQR	2.5	3.0	1.8	3.0	1.0	1.5

Table B.8 shows the descriptive statistics representing the scores of the M-GUDS-S data for the different genders. Yet, the means and medians differ slightly between both groups. Furthermore, homogeneity of variance between the groups for each component of the M-GUDS-S

was assessed with Levene's test. This test yielded no significant results, which means that this assumption was not violated.

Table B.8: The descriptive statistics representing the scores for the M-GUDS-S pre-test and post-test data along the three components for the female group (F) and the male group (M).

	Δ **DC$_F$**	Δ **DC$_M$**	Δ **RA$_F$**	Δ **RA$_M$**	Δ **CD$_F$**	Δ **CD$_M$**
median	0.0	2.0	0.5	0.5	0.5	0.0
mean	0.5	1.6	0.0	1.5	0.2	-1.0
SE mean	0.4	1.0	0.4	1.0	0.4	1.0
sd	2.0	2.7	2.1	2.8	1.9	2.9
IQR	2.3	1.0	3.0	2.8	2.0	0.3

Table B.9 shows the descriptive statistics representing the scores of the M-GUDS-S data for the two groups. Yet, the means and medians differ slightly between both groups. Furthermore, homogeneity of variance between the groups for each component of the M-GUDS-S was assessed with Levene's test. This test yielded no significant results, which means that this assumption was not violated.

Table B.9: The descriptive statistics representing the scores for the M-GUDS-S pre-test and post-test data along the three components for the Dutch group (A) and the abroad group (A).

	Δ **DC$_D$**	Δ **DC$_A$**	Δ **RA$_D$**	Δ **RA$_A$**	Δ **CD$_D$**	Δ **CD$_A$**
median	0.0	1.5	0.0	0.5	0.0	0.5
mean	0.7	1.0	0.0	1.0	0.0	-0.3
SE mean	0.4	0.7	0.4	0.8	0.3	0.9
sd	2.0	2.6	1.8	3.0	1.4	3.2
IQR	2.3	1.8	2.3	2.5	2.0	2.0

B.4 Descriptive statistics GEQ exploratory

Table B.10 shows the descriptive statistics representing the scores of the GEQ data for the female group (F) and the male group (M). Both the medians as well as the means differ a little between the groups. Furthermore, homogeneity of variance between the two groups on the four components of the GEQ was assessed with Levene's test. This test yielded no significant results, which means that this assumption was not violated.

Table B.10: The descriptive statistics representing the scores for the GEQ data along the four components for the female group (F) and the male group (M).

	Immersion$_F$	**Immersion$_M$**	**Flow$_F$**	**Flow$_M$**	**Competence$_F$**	**Competence$_M$**	**Positive affect$_F$**	**Positive affect$_M$**
median	17.5	16.5	13.0	11.5	14.0	13.5	15.0	16.5
mean	17.9	15.3	12.1	13.0	13.2	13.6	15.9	16.1
SE mean	0.6	1.9	0.9	1.3	0.8	1.3	0.6	1.1
sd	3.0	5.3	4.5	3.6	4.0	3.6	2.8	3.0
IQR	3.0	2.5	7.5	3.3	3.3	3.3	3.0	4.0

Table B.11 shows the descriptive statistics representing the scores of the GEQ data for the Dutch group (D) and the Abroad group (A). Both the medians as well as the means are slightly different between the groups. Furthermore, homogeneity of variance between the two groups on the four components of the GEQ was assessed with Levene's test. This test yielded no significant results, which means that this assumption was not violated.

Table B.11: The descriptive statistics representing the scores for the GEQ data along the four components for the Dutch group (D) and the Abroad group (A).

	Immersion$_D$	**Immersion$_A$**	**Flow$_D$**	**Flow$_A$**	**Competence$_D$**	**Competence$_A$**	**Positive affect$_D$**	**Positive affect$_A$**
median	17.0	17.5	11.5	13.0	14.0	14.0	15.0	16.0
mean	17.3	17.3	11.7	13.3	13.2	13.4	15.8	16.3
SE mean	1.0	0.8	0.9	1.2	0.9	1.1	0.7	0.7
sd	4.4	2.6	4.2	4.2	3.9	4.0	3.0	2.5
IQR	4.3	1.5	6.5	6.3	3.3	2.5	4.3	2.5

Appendix C

Initial interviews

C.1 Materials

This protocol, an audio recording device and a notepad to take notes.

C.2 Semi-structured Interview

C.2.1 Opening the interview

Build rapport with the participant so that they feel comfortable being interviewed:

- Hi, what is your name? And how are you?
- Are you ready to start?

C.2.2 Introducing the research

Thank you for joining my study, I appreciate it a lot. For my master thesis, I will design a serious game about diversity and student integration at Utrecht University. Choices will play an important role in the game, as the player's decisions could affect the outcome of the story. The goal of the interview is to receive inspiration for what situations could be in that story. The target group for the game are UU students. Because you are a UU student, I think your answers are very valuable for the design of the game. The interview will take about 15 minutes.

Here is the information sheet and the informed consent form. Take your time to read it carefully. After that, you can decide if you want to sign it or not.

[*if the participant does not sign*] Thank you anyways for your time. Here are the sheets with my contact information in case you change your mind. It is no problem at all to do the interview at a later date. Have a nice day.

C.2.3 Beginning the interview

OK, then the recording starts now. [*start recording*] Let us begin the interview. I will ask some background questions. You do not have to answer if you do not want to disclose this information.

Establish rapport and ask background questions:

- What is your age?
- What is your gender?
- What is your race/ethnicity?
- What is your nationality?
- What do you study at UU?
  - Can you tell me a bit about your study?
  - For how many years have you been studying at UU?

C.2.4 Main part of the interview

Thank you for your answers, we will now move on to the main part of the interview. I just want to emphasise that the questions are not meant like a test, or meant to make you uncomfortable if you do not know what to answer. If anything is unclear, just ask.

I will start with a question about serious games.

- Have you heard of serious games before?
  - If yes, do you know some?
  - *If not, explain briefly what it is*
- Do you know the *Dilemma Game?*
  - If yes, can you tell a bit about it?
  - *If not, explain briefly what it is (similar to the card game used in Dilemmas of the Scientist), and show an example*
  - What do you think of it?
  - Can you tell me what you think the message is?

Since the *Dilemma Game* is focused on diversity, let us dive a little deeper in that topic.

- Have you heard of the terms 'diversity' and 'inclusion' before?
  - If yes, could you tell me what it means?
  - *If not, explain them briefly: While the terms are related, they are not exactly the same. Diversity is about the (in)visible differences between people. Inclusion is the feeling that people feel welcome and valued, despite their differences. So, it is possible to have diversity without inclusion.*
- How important are diversity and inclusion to you?

The *Dilemma Game* shows several scenarios in the context of diversity. Examples are dilemmas about gender and sexuality, different cultures and religions. I am wondering if you also know some examples. These can be from your own experience, or maybe you have heard something from others.

- Could you think of such an example?
  - *For example, a situation where you felt excluded or being 'othered'*
  - *If not, just discuss a few dilemmas from the game and ask for feedback*

Each dilemma also has four options. I plan to use two or three options in the game.

- Could you think of 3 options in the dilemma you just mentioned?
  - *These options can be good, neutral or bad*

Thank you very much for your answers. I have one question left. Now that we discussed some diversity dilemmas and situations, is there any situation *Diverse Perspectives* should absolutely have?

C.2.5 Closing the interview

Wrap up the interview:

- The interview is coming to an end. So we will start to wrap it up. Is there anything else you would like to add?
- Thank you again for participating in the interview. I appreciate it very much. Your data will be anonymised and used as input for the design of *Diverse Perspectives*. Feel free to contact me anytime if you have any questions, remarks or anything else. Have a nice day [*end recording*].

Appendix D

Focus Groups

Table D.1 shows the protocol of the focus group session. Note that in the first part (until *Short break)* a laptop is needed to complete the survey and test the Figma.

Table D.1: Timeline of the focus group session, which should be done in-person. The total estimated time of the timeline is 60 min.

Task	**Materials**	**Time (min)**
Welcome and thank participants, introduce everybody and explain the purpose of the focus group	This document	5
Read information sheet and sign informed consent form, collect demographic data and pre-test M-GUDS-S questionnaire	Laptop, Qualtrics	10
Test the low-fi prototype	Laptop, link to the Figma	10
GEQ and post-test M-GUDS-S questionnaires	Laptop, Qualtrics	5
Short break	Snacks	5
Discuss topic list	Recording device	20
Thank participants and wrap up	A small reward, e.g. a snack	5

D.1 Demographic questionnaire

It's not required to fill in a question if you don't feel comfortable answering.

- What is your age?
- What is your gender?

    - Female
    - Male
    - Non-binary/third-gender
    - Prefer not to say
- What is your nationality?
- What race/ethnicity best describes you?
    - Asian or Pacific Islander
    - Black or African American
    - Hispanic or Latino
    - White or Caucasian
    - Other, namely _____
- Do you identify yourself as a part of a marginalised group?
    - Yes
    - No
    - I don't know/Prefer not to say
- Do you experience discrimination or exclusion based on your own aspects?
    - Yes
    - No
    - I don't know/Prefer not to say
- For how many years have you been studying at UU? Please only use one decimal if these are not whole years, e.g. 1.5 years
- How familiar are you with diversity and inclusion (1 = not at all, 5 = extremely familiar)?
- How familiar are you with serious games (1 = not at all, 5 = extremely familiar)?

D.2 Topic list

- Set-up
    - Pre-test survey
    - Game
    - Post-test survey
- Game itself
    - First impression
    - Perspective switch

- Content of the dilemmas
- Identification with the characters
- Choices align with real life
- Clarity of story
- Fun, engaging
- Design, aesthetics
- Useful, educative
- Other feedback

Appendix E

Experiment

E.1 Materials

This protocol, an audio recording device, a notepad to take notes and some refreshments.

E.2 Semi-structured Interview

E.2.1 Opening the interview

Build rapport with the participant so that they feel comfortable being interviewed:

- Hi, what is your name? And how are you?
- Are you ready to start?

E.2.2 Introducing the research

Thank you for joining my study, I appreciate it a lot. For my master thesis, I designed a serious game about diversity and student inclusion at Utrecht University. Choices play an important role in the game, as the player's decisions could affect the outcome of the story. The goal of the experiment is to research what the impact of *Diverse Perspectives* is on one's attitude towards diversity. The target group for the game are UU students. Because you are a UU student, I think your answers are very valuable for the research. The experiment has two parts. This is part one and will take about 45 minutes [*explain steps of part 1]*. Part two follows after a week, is completely online and takes around 5-10 minutes [*explain steps of part 2*]. Do you have any questions at the moment?

Here is the information sheet and the informed consent form. Take your time to read it carefully. After that, you can decide if you want to sign it or not.

[*if the participant does not sign*] Thank you anyways for your time. Here are the sheets with my contact information in case you change your mind. It is no problem at all to do the interview at a later date. Have a nice day.

E.2.3 Beginning the interview

[*let participant fill in the first part of the survey and play the game, ask if the participant would like a short break after playing the game*] Now that you have played the game, let us

begin the interview. The interview takes about 10-15 minutes. Then the recording starts now [*start recording*].

E.2.4 Main part of the interview

I will start with a question about the game.

- What was your first impression of the game?
  - Fun, engaging, boring, confusing, etc.
  - Which ending did you get?

Thank you for your answer. My next question is:

- What did you think of the visuals?
  - How did you find the layout to navigate through the game?

That's clear. Let's dive deeper into the content of the game.

- What did you think of the dilemmas?
  - And the dialogue flow?
  - (in version A) And the perspective switch?

OK, and regarding the choices you made:

- Would you make the same choices in real life?
  - Why (not)?

To move on to another aspect:

- What did you think of the length of the story?
  - Too long, too short, etc?
  - What do you think of the idea of more episodes?

Thanks. The next question is about the people in the game.

- How much did you feel like you could identify with the characters?
  - And your avatar?
  - Did it feel like you were making your own choices or the avatar's choices?

Alright. I'd like to zoom out, so we're talking about the game in general.

- What did you think about the 'serious' part of the game?
  - Would it be useful in education about diversity?

Good. Then my final question is:

- Do you have any other feedback?

E.2.5 Closing the interview

Wrap up the interview:

- The interview is coming to an end. So we will start to wrap it up. Is there anything else you would like to add?
- Thank you again for participating in the experiment. I appreciate it very much. Your data will be anonymised and used as input for the research of *Diverse Perspectives*. Feel free to contact me anytime if you have any questions, remarks or anything else. Have a nice day [*end recording*].

Appendix F

Shortened Version of the Miville-Guzman Universality-Diversity Scale (M-GUDS-S) questionnaire

Here are the fifteen items of the M-GUDS-S, divided into three subscales. Each item should be answered on a 6-point Likert scale, where 1 = strongly disagree and 6 = strongly agree (Fuertes et al., 2000; Kottke, 2011; Trolian & Parker III, 2022).

Diversity of Contact

1. I would like to join an organisation that emphasises getting to know people from different countries.
2. I would like to go to dances that feature music from other countries.
3. I often listen to music of other cultures.
4. I am interested in learning about the many cultures that have existed in this world.
5. I attend events where I might get to know people from different racial backgrounds.

Relativistic Appreciation

1. Persons with disabilities can teach me things I could not learn elsewhere.
2. I can best understand someone after I get to know how they are both similar and different from me.
3. Knowing how a person differs from me greatly enhances our friendship.
4. In getting to know someone, I like knowing both how they differ from me and is similar to me.
5. Knowing about the different experiences of other people helps me understand my own problems better.

Comfort with Differences

1. Getting to know someone of another race is generally an uncomfortable experience for me.
2. I am only at ease with people of my own race.
3. It is really hard for me to feel close to a person from another race.
4. It is very important that a friend agrees with me on most issues.
5. I often feel irritated by persons of a different race.

Appendix G

Game Experience Questionnaire (GEQ)

Here are the twenty-one items of the first four subscales of the GEQ questionnaire. Each item should be answered on a 5-point Likert scale, where 0 = not at all and 4 = extremely (IJsselsteijn et al., 2013).

Immersion

3. I was interested in the game's story.
12. It was aesthetically pleasing.
18. I felt imaginative.
19. I felt that I could explore things.
27. I found it impressive.
30. It felt like a rich experience.

Flow

5. I was fully occupied with the game.
13. I forgot everything around me.
25. I lost track of time.
28. I was deeply concentrated in the game.
31. I lost connection with the outside world.

Competence

2. I felt skillful.
10. I felt competent.
15. I was good at it.

17. I felt successful.

21. I was fast at reaching the game's targets.

Positive Affect

1. I felt content.

4. I thought it was fun.

6. I felt happy.

14. I felt good.

20. I enjoyed it.

Appendix H

Information Sheet and Informed Consent Form

H.1 Information Sheet

H.1.1 Study Information

This study is part of the Human-Computer Interaction (HCI) master thesis. The thesis researches the design of an inclusive serious game for (international) students at Utrecht University (UU), called *Diverse Perspectives*. A serious game is similar to entertainment games, but the difference is that they do not have entertainment as the main goal. Instead, *Diverse Perspectives* aims to encourage diversity. The game is targeted at UU students.

H.1.2 Purpose [initial interviews]

Diverse Perspectives is currently in the design phase. The purpose of this interview is to obtain inspiration and examples on how diversity could be promoted through the game. In other words, your views and experience are very valuable input for the scenario design of the game.

H.1.3 Purpose [focus groups]

Diverse Perspectives is currently in the design phase. The purpose of this focus group is to obtain feedback on the low-fidelity prototype. In other words, your views and experience are very valuable input for the further design of the game.

H.1.4 Purpose [experiment]

The purpose of this experiment is to research what the impact of *Diverse Perspectives* is on one's attitude towards diversity. Moreover, the experiment aims to investigate which factors contribute to a positive gaming experience.

H.1.5 Your Rights

Any materials produced during the interview may be used for the thesis or other education-related purposes. **During the interview, an audio recording will be taken.** The mate-

rials will be fully anonymised and will be rectified or deleted upon request. **Participation is voluntary.** You can withdraw from the interview at any time without giving any reason. If you withdraw, your materials will be deleted. If you have any questions, ask the researcher or send an email to Geertje Peters Rit (`g.w.f.x.petersrit@students.uu.nl`). In case of any issues, you can contact the thesis supervisor Dr. Hanna Hauptmann (`h.j.hauptmann@uu.nl`).

H.2 Consent Form

In participating in this study, you agree to the following statements below:

- ☐ I confirm that I am at least 18 years of age or older.
- ☐ I confirm that the study has been explained to me. I have been given the opportunity to ask questions about the study and these have been answered sufficiently.
- ☐ I consent that the material I contribute is being used *anonymously* for the HCI master thesis or other education-related purposes.
- ☐ I understand that participation in the study is voluntary and that I can withdraw at any time without giving any reason.
- ☐ I understand that I can request to rectify or delete any data collected from me.
- ☐ I agree to participate in the study.

- - - -

Appendix I

Storyline

The text in blue denotes the additional content of version A that version B does not have.

In this game, you will step into the shoes of a student who is about to begin their very first day at Utrecht University. While you currently stay in a hostel, you're in desperate need for something more permanent. Travelling back and forth is no option, as your hometown is too far away. Finding a room is not easy peasy lemon squeezy...

Characters and situations are purely fictional. Your choices affect the outcome of the story. Choose wisely.

I.1 Scene: At home

[alarm sound beeps] You wake up and look at your phone. It's 8:30 AM, maybe 5 more minutes?

(A) Snooze

(B) Get up

If A: You hit the snooze button... zzz ... But 5 minutes later, the ungodly sound wakes you again.

You stretch your arms and sit up straight in your bed. What day is it? Oh right, it's Monday, which means that the summer break has officially ended. Sad. You look around in your hostel room and ask yourself: "Who do I want to be?"

(A) Female

(B) Male

(C) Non-binary

Then you try to remember more. A question pops up in your head: "What country am I from?"

(A) The Netherlands

(B) Abroad

It all comes back to you now. Your name is

- If A: Jip and today is your first day at Utrecht University. While you have visited several universities on many open days, Utrecht is all you wanted. The city has something magical. While it has canals like Amsterdam, it's smaller and thus more cosy. However, with over 360,000 inhabitants, Utrecht is not small. In 1808, Napoleon's brother made Utrecht even the capital city! However, that didn't last long, as Amsterdam regained that position.

  Although in high school the science courses were not easy, you're eager to start studying Artificial Intelligence. As you're in your early 20s, you also can't wait to live on your own and have your parents out of your hair. It's about time. The only downside of moving to Utrecht will be that you have to leave Stitch behind, your precious cat.

- If B: Sam and today is your first day at Utrecht University. After carefully researching options for studying abroad, Utrecht stood out for you. The city has something magical. While it has canals like Amsterdam, it's smaller and thus more cosy. However, with over 360,000 inhabitants, Utrecht is not small. In 1808, Napoleon's brother made Utrecht even the capital city! However, that didn't last long, as Amsterdam regained that position.

  Although in high school the science courses were not easy, you're eager to start studying Artificial Intelligence. As you're in your early 20s, you also can't wait to live on your own and have your parents out of your hair. It's about time. The only downside of moving to Utrecht will be that you have to leave Stitch behind, your precious cat.

You get out of your bed and open your suitcase for some clothes. Yeah, that outfit will do. As you get dressed and prepare for the day, you think: "What should I eat for breakfast?"

(A) Bread with *hagelslag* (chocolate sprinkles)

(B) Yoghurt with fruit and muesli

(C) Nothing

- If A: Maybe you'll even add a dot of peanut butter for the sweet tooth. You actually never tried it before, but you've heard it's yummy.

- If B: You slice an apple and a banana and put it in a bowl. Then you add the muesli on top. *Smakelijk*!

- If C: Skipping breakfast saves time! But you put a snack in your bag to eat later. *Lion* candy bars are your favourite.

I.2 Travelling

After brushing your teeth and grabbing your backpack, it's time to go to the campus. You go through the door and leave the hostel. As you close your eyes, you think about your day.

While you are glad you have a place to stay for now, you would rather find housing for a longer period of time. Preferable with more space, but less expensive. Luckily, this evening you are invited to a *hospiteeravond*.

You wonder how the evening will play out. Basically, you are invited to a house viewing, which is called a *hospi*, in short. The house is looking for new roommates.

Usually, several people are invited and you have the chance to see the house and the room. A small party can be held, but in the end, it is all about who is the best match for the house.

Or in this case: two matches, as there are two rooms available. *Huize Peereboom* is looking for one Dutch and one international student, to maintain the balance.

Likewise, *Huize Peereboom* finds it important that everyone gets along well, so they prefer Dutch and international students who are already close.

So, it seems to be in your benefit to befriend someone from a different nationality. Also, you could try to interact with strangers too, as both could increase your chances of getting a room. You never know when you will run into your potential roommates...

- If Jip: You already imagine how it would be if you'd live there. Will your new roommates be as excited as you about playing card games? *Ligretto* is your favourite, but you also enjoy playing *Saboteur*.
- If Sam: You already imagine how it would be if you'd live there. Decorating your room with cute plants from that Dutch garden store– what was it called again? *The Inner Time*? *The Intratuin*? Anyway, you're looking forward to it. Your room at home is full with pancake plants.

A notification from your phone brings you back to the present. You should leave now or you'll be late! How will you go to campus?

(A) By bike

(B) By bus

- If A: You find your keys, unlock your bike and hop on it. Just go with the flow. There is no other city in the Netherlands that has as many cyclists as Utrecht. The traffic is quite overwhelming during rush hour and your top priority is not crashing into anyone or anything. Easier said than done... While you are waiting for a red light...
- If B: You grab your OV-card and head to the nearest bus station. The bus you take has the size of three regular buses, being 25 metres long. Woah! Luckily, you find a spot that's free. On the bus...

<u>Cultural differences: Ethnicity</u>

[Dilemma #1] You hear someone joking about ethnicities by using a derogatory term. Chad: "Hehe, of course those *tacoheads* take as much free food samples as possible." When another person confronts him with it, Chad replies: "Whatever. I have friends who belong to this group, so I can do that. They don't seem to mind it." You decide that this is none of your business and try to ignore the conversation, but the person who confronted Chad earlier doesn't buy it. It looks like things are going south... What will you do?

(A) Wait until it passes over

(B) Try to cool it down

- If A: You put in your earbuds and crank up the volume to 100. Bye outside world and its drama.
- If B: You attempt to break it off: "Excuse me, but I would appreciate it if you would stop arguing. It stresses me out, please. Thank you."

Unfortunately, your strategy doesn't work and the two people continue disagreeing. Hence, you try to distract yourself by texting your best friend, describing the situation you're in. They reply with: "OK, yeah that sucks. I would try not to get involved in that, but what's your own opinion about it? Do you think that joke was appropriate?"

(A) (neutral) **I don't know, not enough context.**

(B) (good) **No.**

- If A: "I think I can't be the judge of that. I don't know these people at all."
- If B: "Of course not. People shouldn't make jokes about this and try to come up with lame excuses when others call them out."

You look up from your phone and notice the discussion has ended. At least verbally.

Suddenly, you are feeling weird. You close your eyes, hoping the feeling will go away. Luckily, it vanishes as soon as it came. But when you open your eyes, you're in Chad's body! What's happening? You can hear his thoughts, Chad seems to be continuing the argument in his head:

"Oh geez, people are so sensitive these days! Who do you think you are?! You don't know anything about me or my friends. Back off. You got nothing to say anymore? Then my logic must be solid."

The odd feeling returns. Once you open your eyes, you're back in your own body. Life is... strange.

Chad felt it too.

- [player] "That was weird..."
- [Chad] "Yeah, I noticed it too..."
- [player] "Do you have any idea what happened?"
- [Chad] "No clue... at all."

Before you think about what just happened, Chad and the other person leave and you continue your travel towards the campus. You try to look back on what was taking place just moments ago. Did you just swap bodies for a moment? No, that can't be, as you couldn't control Chad, you could only hear him thinking. Maybe you could use these new mind reading powers to your advantage... for the greater good, right?

I.3 At the UU campus

Finally, you arrived at the Utrecht Science Park, or as many just call it *De Uithof*.

[Dilemma #2]
[if bike]: You bike towards the bike shed. Just when you want to take a turn, someone crosses the street. What do you do?

(A) (neutral) **Dodge them and pretend it didn't happen.**

(B) (bad) **Yell at them and get angry.**

(C) (good) **Stop and ask them where they are going.**

[if bus]: You check out in the bus, put your OV-card away and look around. As you're about to cross the street, someone almost crashes into you. What do you do?

(A) (neutral) **Dodge them and pretend it didn't happen.**

(B) (bad) **Yell at them and get angry.**

(C) (good) **Stop and ask them where they are going.**

You look closely. It's the same person that confronted Chad when he made that joke! The other person recognises you too.

- If A: You wonder if they are headed in the same direction as you.
- If B:
  - [NPC] "It's you again! Watch out!"
  - [player] "Watch out for yourself!"
- If C:
  - [player] "I'm so sorry!" I didn't see you, I swear!"
  - [NPC] "Don't worry about it! I guess we were both not paying attention."
  - [player] "Yeah... where are you going by the way?"
  - [NPC] "Some kind of building called KBG. I'm not sure where it is. [I just arrived in the Netherlands.] Today is my first day on campus."
  - [player] "Me too! I still need to get familiar with Utrecht. It would be easier if I already had a room here, but that's not the case."
  - [NPC] "Same, but hopefully I can move to Utrecht soon. Tonight, I'm invited to this *hospi*, so I might have a chance."
  - [player] "Really? I'm invited to a *hospi* too! Wait, I think it's the same one!"
  - [NPC] "Then you probably also know they prefer one international student and one Dutch student who are already close."
  - [player] "True. But wait, that could be us, right?"

- [NPC] "Yeah, you're right! We can become friends to improve our chances?"
- [player] "Sure! Sounds like a good plan."
- [NPC] "I'm Jip/Sam by the way. And what's your name?"
- [player] "Jip/Sam."

Identity: Transgender and non-binary people

[if befriended] [Dilemma #3A] A group of students passes you. You hear them talking about pronouns. Carmen says: "These non-binary pronouns don't work for me. I find it too much effort, he's not here and I don't like him anyways. Plus, these pronouns don't even exist in my native language. So I'll just stick to using *him*."

- [NPC] "O my god. Did you hear that?"
- [player] "Yes.
- [NPC] "What do you think about what Carmen said?"
- [player]:
  - (A) (good) **I disagree.**
  - (B) (neutral) **No opinion.**
  - (C) (neutral) **I could see why they would see it that way.**
- [NPC] "Yeah, me too."

- If A: Before you know it, [NPC] meddles in the conversation. "You should respect somebody's pronouns. Regardless of their presence and if you like them or not. If you're unsure how to refer to them in your native language, you could ask what their pronouns are."
- If B: You both don't pay further attention to the conversation and wait until they pass by.
- If C: Before you know it, [NPC] meddles in the conversation. "I understand your reasoning. Otherwise you have to invent a word that doesn't even exist in your native language."

The feeling of having a déjà vu pops up. Again, you can hear the thoughts of Carmen.

- If A: "I wasn't talking to you, why are you interfering?!"
- If B: "Nobody objects, so what I said must make sense, right?"
- If C: "Exactly, that's what I mean! Why do people find it such a big deal anyway?"

You blink with your eyes and you're back in your own body again. The telepathic thing still feels weird. Before you can think further about it, Carmen says:

- If A: "I tried asking his pronouns once, but they just don't stick with me. I can't get used to it and I hardly find it the effort, since I rarely see that person."

    - [NPC] "OK, but I think you should give it another shot and ask again. It may not seem a big deal to you, but it can mean a lot to them."
    - [Carmen] "Hmm, maybe you do have a point... I'll think about it."

- If C: "Yep, you read my mind. Look, I'm not trying to be a bad person. In my native language, I'm just used to binary pronouns."

[beep!] Your phone reminds you your class starts within 15 minutes. "Lecture 1: Introduction to AI", it says. You don't want to be late on your first day!

You check your schedule in the UU app and it says your class is in *KBG Cosmos*. You have no idea where that is. You look around and [don't see] Jip/Sam/the other person.

[if befriended]: So, you ask the other person and walk together to class.
[if not]: So, you use your best friend Google Maps to find your way.

After some walking, you arrive at the KBG building. So many windows... In the distance, you also see a sign that says "Botanical Gardens". You put that mentally on your Utrecht bucket list. Perks of being a student is free entrance!
[if befriended]:

- [player] "Hey, have you been to the Botanical Gardens before?"
- [NPC] "Not yet, but I'd like to!"
- [player] "Me too! Want to go together someday?"
- [NPC] "Sounds good! Do you like plants too?"
    - [Jip] "I actually like skateboarding more, but I heard there's a butterfly garden too and I just have to see it. The *Morpho menelaus* is my favourite. It's the blue butterfly."
    - [Sam] "I love plants! Back at home, I have 6 pancake plants in my room. No joke."
    - [NPC] "Cool! [points at the KBG door] Let's go inside."

<u>Identity: Transgender and non-binary people</u>
You go through the door and a sign says the Cosmos lecture hall is on the first floor.

[if alone] [Dilemma #3B] In front of the stairs, someone approaches you. They seem to have some kind of paper in their hand. "Hey, could I ask you something really quick?" "Of course", you say, while hoping they don't want to sell you some kind of newspaper subscription. You have a hard time saying 'No' to those things. "I'm Leo from the Diversity & Inclusion Committee here at Utrecht University. Next month we want to raise awareness for transgender and non-binary people, as they are often misgendered or misunderstood. To estimate how to approach this campaign, I'm gauging the students' opinions about this. Could I ask your opinion about this topic? What do you think of it?"

- (good) **Support it.**
- (neutral) **No opinion.**

- (neutral) **Be sceptical.**
- If A: "I think it's a good idea. I think equipping people with more knowledge about the topic will definitely help. It'll hopefully solve a lot of misunderstandings."
- If B: "To be honest, I don't know, because it doesn't cross my mind often."
- If C:
  - [player] "On the one hand I think it's good, but on the other hand, making people aware of their marginalised positions could have negative consequences, especially if it's not in their power to change it."
  - [Leo] "That's true, I haven't looked at it that way. We have to approach it very carefully, the last thing we want is any backlash."

"Thank you for your answer and your time!" Leo says. "Have a nice day." "You too", you reply.

After climbing the stairs, a pink crocodile catches your eye. Uhm, the what?? Oh, it's the mascot from the *Spar University*! You still have time for a drink. What will you take?

(A) A coffee

(B) A tea

(C) Nothing

- If A/B: Although you find the prices at this 'supermarket' ridiculously overpriced, you enjoy your hot beverage.
  - [if coffee]: How do you like your coffee?
    (a) Just black
    (b) With sugar and cream
  - [if tea]: Which tea flavour do you get?
    (a) Earl grey
    (b) Melon
      - [if befriended and Melon]:
        - [NPC] "Wait, you seriously like melon tea?"
        - [player] "Yes. You don't?"
        - [NPC] "No... why would I drink something that tastes like warm fruit?"
        - [player] "Why not? It's delicious, you should try it."
        - [NPC] "No, thanks..."
- If C: "Way too overpriced", you think. Next time you'll just bring your own thermos.

<u>Disabilities: Physical</u>

[Dilemma #4] While you are walking in the building, you see a girl [member of *Huize Peereboom*] who's in a wheelchair. Matilda's struggling to open the door, but she hasn't asked for any help. What will you do?

(A) (bad) **Pretend you don't see her.**

(B) (neutral) **Immediately help her.**

(C) (good) **Ask if she needs any help.**

- If A: You pretend not to see Matilda's struggle. When she's looking at you, you quickly look in another direction. This isn't your finest moment.
- If B: You rush towards the door to hold it open. Matilda looks at you with slight confusion, but quickly says: "I almost got it, but thanks!"
- If C:
  - [player] "Hey, can I help you with that?"
  - [Matilda] "Actually yes! I didn't want to ask anyone, thank my ego for that. But the door is too heavy."

[if befriended] [Another perspective switch! What is she thinking?]

- If A: "Why are you pretending not to see me? I can clearly see you. And now you look away. How rude."
- If B: "Maybe I could've done it myself, maybe not. It was kind that that person helped me, but I'm not as helpless as I may look."
- If C: "Oh, wow this actually saves me a lot of effort, how nice! I'm really glad you helped me, but asked first. Some people just automatically assume I can't do anything, but that's not the case."

[Within the blink of an eye you're back in your own body.]

If B or C: You wonder why Matilda is in a wheelchair. Will you ask?

- No
- Yes
- If B:
  - [player] "Hey, I hope you don't mind that I ask this, but how did you end up in a wheelchair?"
  - [Matilda] "No, it's fine. I've been in a car accident that resulted in a spinal cord injury. So my legs and feet are paralysed."
  - [player] "Oh... I'm very sorry to hear that."
  - [Matilda] "Thank you, but don't be. It wasn't your fault, right? [laughs] It also happened so long ago... I can handle myself in this wheelchair, but I still miss the things I did before the accident."
  - [player] "What were those?"

- [Matilda] "Well, I used to take long walks through the parks. Utrecht has so many beautiful parks, but the *Máximapark* is my absolute favourite. I can still visit the park in my wheelchair, but it's just not the same."
- [player] "I see... I don't know what else to say..."
- [Matilda] "It's OK, talking about the *Máximapark* brings up some good memories. I'm glad you asked, usually people just see me as kinda helpless and act awkward. I got to go now, but it was nice talking to you. Bye bye!"
- [player] "I have to go too. See you later!"

I.4 In the lecture room

Big, grey letters on the wall spell "Cosmos". That must be the lecture room. You open the door and enter just in time. Quickly, you pick a seat [and Jip/Sam sits next to you].

The professor starts talking: "Welcome everybody to your very first class of *Introduction to AI*. I'm dr. Max Caulfield and I hope you all enjoyed your summer break. Let's start." Dr. Caulfield's monologue continues for a while, but soon your thoughts drift away. You wonder how many international students there are at UU.

[if befriended] Jip/Sam says: "I read that there are more than 35,000 students in total, so I bet around 8,000 of them are international. Wait, let's make it a real bet. If your guess is the closest,

- [if melon tea]: I'll try a cup of that 'tasty' melon tea you had earlier.
- [if not]: I'll buy you a drink.

But if I win, you owe me a drink.". "Deal.", you say.

How many international students are there at UU?

(A) Around 12,000

(B) (correct) Around 6,000

(C) Around 2,000

After consulting your best friend Google, you see that there are almost 6,000 international students. "That's interesting", you think.

[if befriended]:

- [if won and melon tea]:
  - [player] "Ha! Do you want that tea now or...?"
  - [NPC] "I would actually prefer never."
- [if won]:

- [player] "I guess you owe me a drink now."
- [NPC] "Yep, a bet is a bet!"

◦ [if lost]:

- [player] "I guess I owe you a drink now."
- [NPC] "Yep, a bet is a bet!"

After 45 minutes, it's time for a break. You hear some students talk about their plans for going out this evening. "Should I go to *Ekko* or *Tivo*?" a girl says. "What's *Tivo*?", you ask. She replies: "*Tivo* is short for *TivoliVredenburg*. It's a big building with multiple concert halls where you can listen to all kinds of music or go to a party. It's in the middle of the city centre and tonight there is a silent disco. Tickets are €9." "Cool, thanks!", you say. You're excited about discovering Utrecht's nightlife. The more reason to find a room as soon as possible.

Identity: Sexual orientation

[Dilemma #5] The group continues their conversation about their evening plans. Some students are talking about going to café *BodyTalk*. "What is that?" you ask. "It's a LGBTQ+-friendly café.", Chloe replies. "I went there with my girlfriend last time." Suddenly, another student swamps Chloe with questions: Are you gay? Since when? How did you come out? How can you be so sure? Did you ever have a crush on one of your friends? The conversation abruptly seems to have turned into an interrogation. You can see that Chloe doesn't like it and shuts down. What do you do?

(A) (neutral) **Switch the topic.**

(B) (good) **Say something about it.**

(C) (neutral) **Wait until the topic switches, but talk later in private.**

- If A: You try to change the topic. "Sooo, nice weather we're having, right?"
- If B: You let the person know that they are making Chloe uncomfortable with their questions and suggest they should stop. You see that the girl is visibly relieved.
- If C: You wait until the topic switches and talk to Chloe privately how she feels about it. "I got myself back together and I'm OK now. Sorry you had to see my inner meltdown." she replies. "I just wish it wasn't such a big deal every time I brought it up." "Don't be sorry", you say. "The one that should be, is that person who started those invasive and inappropriate questions." Chloe: "You're right."

[if befriended] [Before you notice, you can read the other's mind again.]

- If A: "OK, I guess that awkward attempt kinda worked? But I still can't shake the feeling of harassment, this way people will never realise how invasive and inappropriate their questions actually are. You wouldn't ask a straight person the same questions, would you?"
- If B: "Wow, I'm really glad you put them in their place. People are sometimes just so nosey. You wouldn't ask a straight person the same questions, would you?"

- If C: "Thanks for checking in, but I wish you'd done that earlier. I mean, why didn't anybody say anything? I feel so alone, this makes me really sad. I probably shouldn't have mentioned it, but it still feels unfair. You wouldn't ask a straight person the same questions, would you?"

[Nevermind...]

- If A: The group looks confused at you. However, they all sense the tense atmosphere, so they're happy to switch the topic.
- If B:
  - [Chloe] "Thanks so much for standing up for me! I'm so done with the fact that it's such a big deal every time I bring it up. As if my whole personality is reduced to having a girlfriend instead of a boyfriend and I owe everyone an explanation."
  - [player] "Yeah, no problem! Glad I could help."
  - [Chloe]: "Are you also studying Artificial Intelligence?"
  - [player] "Yes, I started today!"
  - [Chloe] "Cool, I tried Computing Science before, but that wasn't really my thing. I hope this study will suit me better."

player : "Maybe it will! It's good that you try another study. I wouldn't know what else I would want to study besides AI."

  - [Chloe] "You really don't?"
  - [player] "Hmm, thinking about it, maybe Gender Studies? But I never thought about it much further, I took a lot of science courses in high school instead of the humanities."
  - [Chloe] "Sooo... what are you doing after class?"
  - [player] "I'm going to the city centre. What about you?"
  - [Chloe] "Nothing much. But we can stay in touch if you'd like? Can I get your number? Since you helped me, I want to take you out for a drink."
  - [player]
    (A) Yes. → Here it is. Maybe your girlfriend could join too? I would like to meet her.
    (B) No. → Sorry, but I don't know you that well. But thanks for offering.

Before you can continue the conversation, the break ends. Too soon, as always. The professor begins with the second part of the lecture. Although you find it interesting, your brain is still in holiday mode, which makes it hard to stay focused for more than 10 minutes. Only 45 more minutes left. This is roughly 4 x 10 minutes. Should be doable.

After Max Caulfield explains the group project, it's time to form groups of three and you are free to form them yourselves.

[if befriended]: You team up with Jip/Sam. It's nice to have at least one familiar face in your group.

Formal education

[Dilemma #6] Someone approaches you. Val, a student from the *HU* (applied sciences), asks if they could join your group. What is your response?

(A) (bad) **Sorry, but no.**

(B) (neutral) **Let me think about it.**

(C) (good) **Sure!**

- If A: You refuse politely and you decide to stretch the truth a little: "Sorry, but our group is already full." "OK." Val replies. "I understand."
- If B: You are not sure yet. "Wait, let me check with this other person real quick." "OK." Val replies. "Please let me know later."
- If C: "Of course!" The more diverse backgrounds, the better. You could probably learn a thing or two. "Thank you!" Val replies and they smile.

[if befriended] [Poof! Now you're in Val's head!]

- If A: "Actually, I don't understand it at all. I can't shake the feeling that people look down on me just because I didn't do university..."
- If B: "Maybe I shouldn't have mentioned my previous education. It seems to put me at a disadvantage."
- If C: "That went well. I'm glad I found a group so quickly."

[Reverse poof! Now you're back to being your old self.]

[if no perspective switch: a short moment later]

[Dilemma #6.2]

- If B: Val approaches you and says: "Hey, so have you made up your mind? Can I join your group?"

  (A) (good) **Sure.** → "Of course!" The more diverse backgrounds, the better. You could probably learn a thing or two. "Thank you!" Val replies and they smile.

  (B) (bad) **Sorry, but no.** → You refuse politely and you decide to stretch the truth a little: "Sorry, but our group is already full." "OK." Val replies. "I understand."

[another chance to befriend the NPC again and find out about the *hospi* if the player has accepted Val *and* didn't get angry in Dilemma #2]

- ◦ If A (continued): Another student approaches you:
  - [NPC] "Can I also join your group? [I just arrived in the Netherlands.] It's my first day on campus and I don't know anybody yet."

- [player] "Of course! And it's my first day too! I still need to get familiar with Utrecht. It would be easier if I already had a room here, but that's not the case."
- [NPC] "Same, but hopefully I can move to Utrecht soon. Tonight, I'm invited to this *hospi*, so I might have a chance."
- [player] "Really? I'm invited to a *hospi* too! Wait, I think it's the same one!"
- [NPC] "Then you probably also know they prefer one international student and one Dutch student who are already close."
- [player] "True. But wait, that could be us, right?"
- [NPC] "Yeah, you're right! We can become friends to improve our chances?"

player "Sure! Sounds like a good plan."

- [NPC] "I'm Jip/Sam by the way. And what's your name?"
- [player] "Jip/Sam."

I.5 In KBG

Class has ended and you want to explore the city centre. [if befriended: you'll go together]

- If Jip: You'd also like to visit the skate ramp in the *Griftpark*. The pictures on the internet looked really cool. But you don't have your board with you right now. Maybe later this week?
- If Sam: You'd also like to visit this mysterious shop called *BLACKFISH*. The pictures on the internet looked really cool and you heard that they give away free stickers. But they're closed today. Maybe later this week?

As you walk out of the lecture room, you see that the exit of the building is on the ground floor. Will you take the elevator or the stairs?

(A) Elevator

(B) Stairs

- If A: You take it easy and wait for the elevator to arrive. It doesn't matter that it's just for one floor.
- If B: You swiftly descend the stairs. That was a piece of cake for a young god like you.

While you are heading outside you hear a few people singing a birthday song. It seems like it's the birthday of an Asian student, Shiro. After people finish the classical *Happy Birthday* song, they continue with *Hanky Panky Shanghai*. This song is sung in the same melody as *Happy Birthday* and often people tend to narrow their eyes with their fingers when singing it.

<u>Cultural differences: Ethnicity</u>

[Dilemma #7] It looks like the majority enjoys singing the song. What do you do?

(A) (bad) **Join the group.**

(B) (good) **Ask in private if Shiro is OK.**

(C) (neutral) **Walk past it.**

- If A: You join the group with your excellent singing skills.
- If B: You ask Shiro if he's OK with it. "To be honest, no. Not at all." he replies. "Thank you for asking though. I realise now that I just should tell my friends to stop because it makes me uncomfortable. It's so weird that it is taken for granted that people think it's appropriate to do it just because someone is Asian... But besides that, my friends are actually nice people, although just a bit oblivious."
- If C: You find birthday songs silly anyway and walk away, while giving a quick side eye.

[if befriended][Oh, there's the switch! What are they thinking this time?]

- If A: "O my god why are you joining too? As if this situation couldn't get worse..."
- If B: "I'm so glad that you asked. This is just the boost I needed to put an end to this nonsense."
- If C: "Nobody comments on this... Maybe they don't realise that their 'good' intentions come out really bad."

[And now you're back to normal.]

[Dilemma #7.2] [only if you had the perspective switch]

- If C: Now you start to feel bad for Shiro. Although you don't know him or his friends at all, maybe it's worth a try to say something. What will you do?

  (A) (neutral) **Brush off the feeling and start walking faster.** → Sorry, but you can't take the whole world's weight on your shoulders.

  (B) **Ask if Shiro is OK.** → "To be honest, no. Not at all." he replies. "Thank you for asking though. I realise now that I just should tell my friends to stop because it makes me uncomfortable. It's so weird that it is taken for granted that people think it's appropriate to do it just because someone is Asian..."

Speaking of birthdays,

- If Jip: you suddenly think back about this year's Kingsday. Eating a *tompouce* in the sunshine, visiting the flea market with your friends... seems like ages ago.
- If Sam: you suddenly remember something about a celebration called "Kingsday". And the colour orange. Loooots of orange. Why are Dutch people so obsessed with it?

[if befriended]:

- [Sam] "Hey, can I ask you a question?"

- [Jip] "Of course!"
- [Sam] "Why do Dutch people have a thing for orange stuff? And especially during some kind of celebration day? I don't get it."
- [Jip] [laughs] "Oh, that's because on the 27th of April, we celebrate the King's birthday. And the word 'orange' is actually part of the surname from the royal bloodline, so that's why everything is orange."
- [Sam] "Ah, makes sense."
- [Jip] "In Utrecht, they take Kingsday very seriously. I mean, the celebrations already start at 6 PM the evening before the 27th and take exactly 24 hours. I recommend visiting the flea market, people sell so much stuff and sometimes you find hidden gems. Oh, and eating a *tompouce*, of course!"
- [Sam] "Sounds good, thanks for the tips!"
- [Jip] "You're welcome!"

I.6 In the city centre

As you leave the KBG, you look for your bike/the bus stop.

- If you took the bike earlier: you bike towards the city centre and park your bike in one of the sheds underground. The shed is guarded and the first 24 hours are free. Plus, you don't risk that your bike will be taken away by the municipality.
- If you took the bus earlier: you take bus 28 and get off at the stop called 'Neude'.

You walk along the *Oudegracht*, which is a street in the heart of the centre. The length of the street is almost 2 km! It's also the street where the tiniest house in Utrecht is located: at number 133. At the other side of the street is the oldest house in Utrecht: at number 114. In fact, *Drakenborch* used to be a castle.

From the *Oudegracht*, you could already see the Dom Tower. Impressive!

After a few minutes of walking, you find yourself at the Dom Square, looking at the Dom Tower. You are wondering how tall the Dom actually is.

- [if won bet before]
  - [player] "Wanna bet again?"
  - [NPC] "I'll skip this one."
- [if lost bet before]
  - [NPC] "Wanna bet again?"
  - [player] "I'll skip this one."

How tall is the Dom Tower?

(A) 56 metres

(B) (correct) 112 metres

(C) 224 metres

If A or C: A quick Google search reveals that it's actually 112 metres. Oops.

Neurodivergence

[if B, bonus dilemma] You notice that somebody is standing next to you.

- [Steph] "In case you're wondering, the Dom is actually 112 metres and 32 centimetres! To get to the top, you have to take 465 steps. People even organise stair climbing matches. I like random facts, that's why I know. People find me a bit peculiar, and they might be right. I'm Steph, by the way."
- [player] "I'm Jip/Sam. And why is that so?"
- [Steph] "Just between me and you, I'm diagnosed with autism, but I don't know if I want to tell my friends about this. I'm afraid I will be judged differently when I do. I don't mind telling strangers though." What would you say to Steph?

  (A) (neutral) **Play it down.**
  (B) (good) **Thank and support.**

- If A: "Everybody has some autistic traits and that it's not a big deal in telling people." "You're probably right", Steph replies. "But I'm still not sure if I'm ready for that."
- If B: "I can't make that choice for you, but I support you in whatever decision you take. Thanks a lot for sharing that with me." "Thank you", Steph replies.

[if befriended][Oh, it's happening again! What's on their mind?]

- If A: "Why did I ask a random stranger for advice? Besides that, you probably think I'm blowing it up, but that's not the case. I'm really struggling with it."
- If B: "They're right, it's my own decision to make. I guess I just needed to feel understood."

[Back to hearing only your own thoughts.]

Steph smiles and walks away.

You've also read on the internet that Utrecht is the city where Miffy was born. Or in Dutch: *Nijntje*. You want to visit Miffy's statue at the *Nijntje pleintje* (Miffy square). As you're walking, you cross the *Janskerkhof*. This place has the oldest bookstore in the Netherlands, because the first Dutch book was printed in Utrecht. Moreover, every Saturday, the parking space at the *Janskerkhof* turns into a flower market.

After visiting the Miffy square, you're a bit tired, so you decide to have a drink to sit down and rest. There's a free spot on the terrace. Because of the many canals in the centre, a lot of cafés still receive their beer by boat. In the middle ages, the beer was even brewed by using the water from the canals. Ew! After a few minutes, a waiter comes to your table. "What would you like to drink?"

(A) A beer

(B) A wine

(C) A soda

"A [player's choice] please!" you tell the waiter. [if befriended] "For me too, please!"

A few moments later, the waiter arrives with [player's choice]. "Here you go. Enjoy!"

[if befriended]: "Cheers!" you say and you both toast the glasses. As you are both enjoying your drinks, you're chatting about all kinds of stuff.

- [NPC] "What's your favourite colour?"
- [player]:
  (A) Blue
  (B) Green
  (C) Red
- "It's [player's choice]."
- [NPC] "I thought it would be orange... Hahaha, no I'm kidding! Mine's cerulean."
- [player] "Uhm, isn't that some kind of fancy word for 'blue'?"
- [NPC] "Ouch, no. It's just not the same."
- [player] "Are you serious? It has the same colour as blue curaçao. And they don't call it 'cerulean curaçao' either."
- [NPC] "Fair. And that does have a nice ring to it."
- [player] [laughs] "Hey, don't steal my idea!"
- [NPC] "I would never... [laughs]"

Cultural differences: Language

[Dilemma #8A] [if befriended]: The topic changes and you talk more about the *hospi*. Jip/Sam says that they heard a story from a friend who joined a student house as the only international student. The house also has a group chat where they chat in Dutch. After the international student was added to the group chat, a few roommates were unsure whether to continue talking in Dutch or switch to English. What would you do?

(A) (neutral) **Continue in Dutch, highlights in English.**

(B) (bad) **Stick to Dutch.**

(C) (good) **Switch to English.**

- If A: "I would continue casual communication in Dutch, but for important announcements I would switch to English."

- If B: You say: "As I see it, the majority of the house is Dutch. So I would stick to Dutch."
- If C: "I would talk in English," you say. "So the international student can read the chat too."

[NPC]: "OK, fair. Honestly, I wouldn't know what I would do. I think it also depends how close the house is and if they're used to speaking English in their daily lives."

[no perspective switch, because there's only the two of you]

Cultural differences: Language

[Dilemma #8B] [if you're sitting alone]. As you are enjoying your drink, you see some members of the study association *Sticky* are sitting at a table next to you. They're busy chatting. One member is asked to make an announcement at the next activity. They're saying that they would rather make the announcement in Dutch, since they're much more comfortable speaking Dutch. However, *Sticky* has several international members. If you were in their place, what would you do?

(A) (neutral) **English summary.**

(B) (bad) **Speak Dutch.**

(C) (good) **Speak English.**

- If A: You would perform the announcement in Dutch and give a highlighted summary in English at the end. That seems fair, right?
- If B: You would perform the announcement in Dutch. That way you can express yourself better and that's only fair, right?
- If C: You would perform the announcement in English. While it costs a little more effort from your side, it's more important that everybody receives the message loud and clear.

After you finish your drink, your stomach starts rumbling. Time for dinner! But what will you eat?

(A) Pasta pesto

(B) A *Knorr Wereldgerecht* (random dish)

(C) *Stamppot* (mashed potatoes and veggies, with gravy)

The pro of this plan is dinner, but the con is making it. But it's at least something you like.

- If Jip: No way you will make something that has carrots in it. It's the only thing you despise. But at home, you have to eat what's on the menu.
- If Sam: No way you will make something that has goat cheese in it. It's the only thing you despise. But at home, you have to eat what's on the menu.

Because you still need to buy groceries, you open Google Maps to find a supermarket. Plenty of them in the *Amsterdamsestraatweg*! This street was built on Napoleon's orders in 1812 and is part of the route between Paris and Amsterdam. Since it's the longest shopping street in the Netherlands, finding a supermarket there shouldn't be too difficult. After you pay, you start walking.

I.7 In a supermarket

At the *Amsterdamsestraatweg*, you spot a *Jumbo*. Although not as cheap as *Aldi* or *Lidl*, it's good enough. Standing in the *Jumbo*, you put all the ingredients for a [player's choice] in your basket.

Cultural differences: Religion

[Dilemma #9] [if befriended]: You ask if Jip/Sam wants to eat together. They would like to! As long as there are no carrots/goat cheese in it. And they mention that they only eat halal meat due to religious reasons. You're both at the supermarket buying ingredients for dinner. However, the halal meat is out of stock. What now?

(A) (good) **Go to another store to buy it.**

(B) (bad) **Buy regular meat, but pretend it's halal.**

(C) (neutral) **Just skip the meat.**

- If A: Your friend deserves a tasty meal too. You take the extra mile to buy halal meat at another store.
- If B: You don't take your friend's wishes into account. But they don't need to know that.
- If C: You decide to make the dish with less ingredients. If you don't buy any ingredients that are not halal, they can't cheat, right?

[no perspective switch, because there's only the two of you]

After paying for your groceries, you go home and eat dinner [together].

[if befriended]:

- [NPC] "Thanks for inviting me to your place! I'm staying in a hostel too currently."
- [player] "Yeah, no problem! Wait, in which hostel are you staying?"
- [NPC] "At the *Student Hostel* close to the city theatre. It's actually not so far from the *Stayokay* here at Neude. By the way, can I complain about somebody to you? "
- [player] "Of course! Spill the tea."

Dietary wishes

[Dilemma #10] [if befriended]: During dinner, Jip/Sam starts complaining about a vegan friend. They say that they just don't understand why anybody would be vegan, it feels like they do it to guilt-trip people that do eat meat and dairy products. In other words, they just can't wrap their heads around it and continue their monologue for a while. What do you do?

(A) (neutral) **Say nothing.**

(B) (bad) **Pretend to agree.**

(C) (good) **Interrupt them by disagreeing.**

- If A: Although you don't think the same about it, you wait until their rant is over. While tempting, you don't want to point them out that they have their dietary wishes too.
- If B: You pretend to agree with them and also do your two cents by saying it's only for the elite and pretentious ones. Yet in reality, you don't think that at all, but you don't want a confrontation.
- If C: You interrupt them and explain the reasons why people could be vegan. Your friend shouldn't criticise them for that, that's unfair.

If B [Dilemma #10.2]:

- [NPC] "Glad you think the same about it. But you don't have to agree just because we are friends."
- [player] "I know..."

  (A) (neutral) **Keep quiet.**

  (B) (good) **Speak your mind.**

  - If A: Although you don't think the same about it, you decide to drop it and finish your plate.
  - If B: "Actually, I do disagree. People have their own reasons to be vegan. I think it's unfair you criticise them for that." → go to C.

If C:

- [NPC] "Well, I still agree to disagree. But that's what we're friends for."
- [player] "Yep."

When you finish eating, it's almost time to go to the *hospi*! You give yourself a pep talk and open the door.

I.8 At the *hospi*

[if befriended]: You go to the *hospi* together. This also makes you less nervous, which gives you a good first impression at *Huize Peereboom*.

[if alone]: You arrive at the house. You can't help it, but you're a little nervous. Or a lot. As you walk through the front door, you see the person you've run into earlier today.

- If at least 6 points: the NPC greets you kindly and you immediately feel more confident → + 1 point
- If at least 3 points: the NPC quickly greets you, but starts talking to another person
- If less than 3 points: the NPC ignores you, which makes you even more nervous. This doesn't help you to make a good first impression → - 1 point.

You enter the living room and take a seat. There are quite some people and Matilda is here too! Wait, she lives here?? You remember the encounter you had with your potential new roommate earlier today...

[if befriended] You sit next to Jip/Sam.

[if alone] You sit next to a familiar face.

[good ending, at least 8 points] During the day, you've become close with Jip/Sam. You feel comfortable at the *hospi*. *Huize Peereboom* likes this. At the end of the evening, you get a phone call that you both will be the new roommates! That means no more expensive hostel rooms. And you already made a friend on your first day. This day couldn't get any better!

[neutral ending, at least 4 points] You chat a little with Jip/Sam and do your best at the *hospi*. Sadly, neither of you is picked by *Huize Peereboom* to be the next roommates. At the end of the evening, you both return to your hostel rooms. Maybe better luck next time.

[bad ending, less than 4 points] The other person acts coldly towards you at the *hospi*. This negatively affects you. You don't feel at ease and are therefore not chosen to be the next roommate at *Huize Peereboom*. Instead, you find out that the other person was chosen. You go back to your hostel room and cry yourself to sleep. Hopefully, not all days will be like this...

You close your eyes and fall asleep.

I.9 Recap

You've reached one of the three endings. Congratulations for completing the game! During the story, there were several dilemmas. Most of the dilemmas covered a specific diversity (sub)theme. Here is a chronological overview of all the (sub)themes that you have possibly encountered:

- Cultural differences: Ethnicity

- Identity: Transgender and non-binary people
- Disabilities: Physical
- Identity: Sexual orientation
- Formal education
- Neurodivergence
- Cultural differences: Language
- Cultural differences: Religion
- Dietary wishes

Regarding diversity, this is still the tip of the iceberg. We are all different. Although things can be unfamiliar and scary, try to look at them from a diverse perspective. And when you are not sure about something: just ask instead of assume!

The end. Thank you for playing!